

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-243256

(P2001-243256A)

(43) 公開日 平成13年9月7日(2001.9.7)

(51) Int.Cl. ⁷	識別記号	F I	デマコト [*] (参考)
G 0 6 F 17/30	3 4 0	G 0 6 F 17/30	3 4 0 A 5 B 0 4 9
	1 7 0		1 7 0 Z 5 B 0 7 5
	2 1 0		2 1 0 A
13/00	5 4 0	13/00	5 4 0 P
17/60	3 2 6	17/60	3 2 6
審査請求 未請求 請求項の数16 O L (全 36 頁)			

(21) 出願番号 特願2000-347730(P2000-347730)

(22) 出願日 平成12年11月15日(2000.11.15)

(31) 優先権主張番号 4 8 3 0 9 2

(32) 優先日 平成12年1月14日(2000.1.14)

(33) 優先権主張国 米国 (U S)

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 ジャメイ グラハム

アメリカ合衆国, カリフォルニア 94025,

メンロ・パーク, サンド・ヒル・ロード

2882番, スイート115, リコーコーポレ

ーション内

(74) 代理人 100070150

弁理士 伊東 忠彦 (外1名)

最終頁に続く

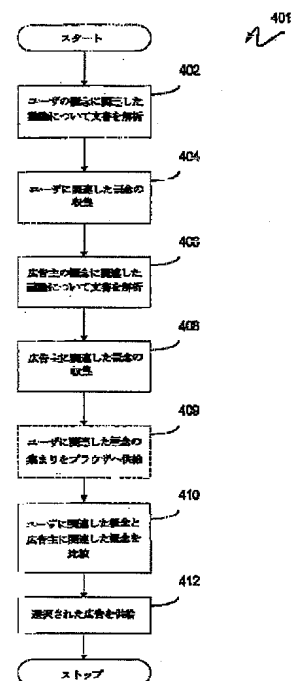
(54) 【発明の名称】 ウェブ広告に基づく内容表示方法及び装置と、内容表示プログラム

(57) 【要約】

【課題】 本発明は、インターネット上で閲覧者の関心に基づいてオンライン広告を閲覧者に配布する目的を絞った販売システムの提供を目的とする。

【解決手段】 本発明の一実施例は、ユーザの関心の概念と、広告主の関心の概念と、閲覧中の文書のマッチングを用いて、文書の閲覧者へ広告を的確に送る。ブラウザで閲覧されているページ毎に前後関係に対応した広告を生成し、広告を文書のページと関連付ける。本発明によりの的確に提供される広告は、たとえば、実質的に無料で広告が組み込まれる文書に関連付けられ、或いは、組み込まれた広告でもよい。

本発明の具体的な一実施例における無道ステップのフローチャート



【特許請求の範囲】

【請求項1】 ユーザへ提示された少なくとも一つの文書の内容に基づいて広告をユーザへの確に提供する方法であって、

ユーザ概念関連性を得るため、ユーザが選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した情報の議論を識別すべく上記少なくとも一つの文書を解析する工程と、

広告主概念関連性を得るため、広告主が選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した情報の議論を識別すべく上記少なくとも一つの文書を解析する工程と、

全体的な関連性を判定するため、上記ユーザ概念関連性と上記広告主概念関連性を比較する工程と、

特定の全体的な関連性を有する広告を選択する工程と、を含む方法。

【請求項2】 上記選択された広告をウェブブラウザの販売情報領域に表示する工程を更に含む、請求項1記載の方法。

【請求項3】 上記比較する工程は、第1の用語の組と第2の用語の組の間の類似性を判定する工程を含み、上記第1の用語の組はユーザ概念であり、上記第2の用語の組は広告主概念である、請求項1記載の方法。

【請求項4】 上記比較する工程は、ユーザ概念 c_i により構成され m 個のユーザ概念を含む第1のベクトル C と、広告主概念 a_j により構成され n 個の広告主概念を含む第2のベクトル A の間で用語の頻度を測定する工程と、

上記第1のベクトルと上記第2のベクトルに共通した概念の一致の指標を含む第3のベクトル $V_{i,j}$ を得るため、上記第1のベクトルと上記第2のベクトルを比較する工程とを有する、請求項1記載の方法。

【請求項5】 上記第3のベクトル $V_{i,j}$ 中の一致の指標 $W_{i,j}$ の個数を判定する工程と、

上記第3のベクトル $V_{i,j}$ の長さ X_i を判定する工程と、

値 c_i を決める工程と、

以下の関係式、

【数1】

$$S_{i,j} = \left(\frac{W_{i,j}}{X_i} \right) \times c_i$$

にしたがって、類似性値 $S_{i,j}$ を判定する工程とを更に有する請求項4記載の方法。

【請求項6】 以下の関係式、

【数2】

$$R_j = \left(\sum_i S_{i,j} \right) \times a_j$$

にしたがって、上記関連性 R_j を判定する工程を更に有する請求項5記載の方法。

【請求項7】 ユーザへ提示された少なくとも一つの文書の内容に基づいて広告をユーザへの確に提供するシステムであって、

ユーザが選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した上記文書に含まれる概念の出力を得るため、ユーザが選択可能な上記複数の関心の概念の中の上記少なくとも一つの概念に対応した情報について、上記文書の内容を解析するプロファイル内容理解器と、

広告主が選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した上記文書に含まれる概念の出力を得るため、広告主が選択可能な上記複数の関心の概念の中の上記少なくとも一つの概念に対応した情報について、上記文書の内容を解析する広告内容理解器と、

ユーザが選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した上記文書に含まれる上記概念の出力を、広告主が選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した上記文書に含まれる上記概念の出力と比較し、蓄積された複数の広告の中から、比較の結果に基づいて上記少なくとも一つの文書に含まれる情報に関連した広告を選択する比較器と、を具備したシステム。

【請求項8】 上記広告内容理解器は、上記蓄積された広告から最良の広告を選択する、請求項7記載のシステム。

【請求項9】 コンピュータに、

ユーザ概念関連性を得るため、ユーザが選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した情報の議論を識別すべく少なくとも一つの文書を解析する手順と、

広告主概念関連性を得るため、広告主が選択可能な複数の関心の概念の中の少なくとも一つの概念に対応した情報の議論を識別すべく上記少なくとも一つの文書を解析する手順と、

全体的な関連性を判定するため、上記ユーザ概念関連性と上記広告主概念関連性を比較する手順と、

特定の全体的な関連性を有する広告を選択する手順と、を実行させ、ユーザへ提示された少なくとも一つの文書の内容に基づいて広告をユーザへの確に提供するためのプログラム。

【請求項10】 第1の関心の概念に対応した情報を識別するため、少なくとも一つのウェブ対象を解析する工程と、

第2の関心の概念に対応した情報を識別するため、上記少なくとも一つのウェブ対象を解析する工程と、

上記第1の関心の概念と上記第2の関心の概念の間の関連性を判定する工程と、

上記関連性に基づいて上記少なくとも一つのウェブ対象を少なくとも一つの広告対象と関連付ける工程とを有する、広告対象をウェブ対象と自動的に関連付ける方法。

【請求項11】 上記関連性を判定する工程は、第1の用語の組と第2の用語の組の間の類似性を判定する工程を含み、

上記第1の用語の組は上記第1の関心の概念であり、上記第2の用語の組は上記第2の関心の概念である、請求項10記載の方法。

【請求項12】 上記関連性を判定する工程は、上記第1の関心の概念 c_i により構成され m 個の概念を含む第1のベクトル C と、上記第2の関心の概念 a_j により構成され n 個の概念を含む第2のベクトル A の間で用語の頻度を測定する工程と、

上記第1のベクトルと上記第2のベクトルに共通した概念の一致の指標を含む第3のベクトル $V_{i,j}$ を得るため、上記第1のベクトルと上記第2のベクトルを比較する工程とを有する、請求項10記載の方法。

【請求項13】 上記第3のベクトル $V_{i,j}$ 中の一致の指標 $W_{i,j}$ の個数を判定する工程と、上記第3のベクトル $V_{i,j}$ の長さ X_i を判定する工程と、

値 c_i を決める工程と、
以下の関係式、

【数3】

$$S_{i,j} = \left(\frac{W_{ij}}{X_i} \right) \times c_i$$

にしたがって、類似性値 $S_{i,j}$ を判定する工程とを更に有する請求項12記載の方法。

【請求項14】 以下の関係式、

【数4】

$$R_j = \left(\sum_i S_{ij} \right) \times a_j$$

にしたがって、上記関連性 R_j を判定する工程を更に有する請求項13記載の方法。

【請求項15】 上記広告対象をウェブブラウザの販売情報領域に表示する工程を更に有する、請求項14記載の方法。

【請求項16】 コンピュータに、
第1の関心の概念に対応した情報を識別するため、少なくとも一つのウェブ対象を解析する手順と、
第2の関心の概念に対応した情報を識別するため、上記少なくとも一つのウェブ対象を解析する手順と、
上記第1の関心の概念と上記第2の関心の概念の間の関連性を判定する手順と、
上記関連性に基づいて上記少なくとも一つのウェブ対象を少なくとも一つの広告対象と関連付ける手順と、を実行させ、広告対象をウェブ対象と自動的に関連付けるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、オンライン広告管理に係り、特に、インターネット環境において、広告情報やその他の販売情報などの情報を的確に導く技術に関する。

【0002】

【従来の技術】電子商取引の急速な発展と共に、ウェブ広告のための顧客ベースの作成を目標とする無料サービスを提供するインターネット企業数は、劇的に増加している。典型的に、インターネット企業は、ポータル、すなわち、多数の他のウェブページへのリンクを含むウェブページを使用するウェブページを運営する。ポータルは、ウェブ広告を含む。ポータル上に広告を出す企業は、ポータルプロバイダに使用料を支払う。普及しているポータルの例には、MYAHOO、MYEXCITE、NETSCAPE NETCENTER、及び、HOTMAILが含まれる。

【0003】

【発明が解決しようとする課題】ウェブポータルを使用する利点は認められているが、より良い効率を得ることができる。たとえば、ウェブポータルは、ユーザがポータルによって制御されるウェブページの領域を出ると、広告をユーザに提示することができない。America On Line(AOL)のようなインターネットサービスプロバイダ(ISP)は、ユーザへサービスされたページを制御し続けるので、加入者へ到達する広告をより多く制御する。しかし、ユーザへ提示された全てのウェブページ上に広告を出し続けるためには、従来、ウェブページをフォーマットし直すか、或いは、広告が掲載された煩わしいウィンドウをポップアップさせる必要がある。

【0004】したがって、本発明は、オンラインユーザのプライバシーを漏洩することなく、広告をオンラインユーザへ的確に導く技術の提供を目的とする。

【0005】

【課題を解決するための手段】本発明によれば、閲覧者の関心に基づいてオンライン広告を閲覧者に配布するインターネットターゲット販売システム、方法、及び、プログラムが得られる。

【0006】本発明の具体的な実施例は、的確な広告を現在文書の閲覧者へ送るため、ユーザの関心の関心の概念と、広告主の概念と、現在閲覧されている文書に含まれる概念の種々の(n-way)マッチングを使用する。

【0007】一部の実施例は、ブラウザ内で閲覧される各ページに前後関係に敏感な広告を生成することができ、これにより、広告を文書内の全てのページと関連付ける。

【0008】特定の実施例は、広告を、たとえば、実質的に無料で広告が埋め込まれる文書と関連付けることができる。

【0009】本発明による具体的な一実施例によれば、使用中のアクティブ文書の内容に基づいて、的確な広告をユーザに送る方法が提供される。使用中の文書とは、

たとえば、現在、ユーザに表示されている文書である。しかし、一部の実施例では、使用中の文書は、デーモン若しくはバックグラウンドプロセスによって処理されている文書である。この方法は、一つ以上のユーザが選択可能な関心の概念に対応した情報の議論を識別するため、少なくとも一つの文書を分析する。この分析によって、ユーザ概念関連性、すなわち、ユーザによって定義された一つ以上の概念にする文書の関連性の尺度が得られる。また、この方法は、一つ以上の広告主が選択可能な関心事項に対応した情報の議論を識別するため、文書を分析する。この分析によって、広告主概念関連性、すなわち、広告主によって定義された一つ以上の概念にする文書の関連性の尺度が得られる。さらに、この方法は、全体的な関連性を判定するため、文書に対するユーザ概念関連性と広告主概念関連性を比較する。特定の全体的な関連性を有する広告の選択は、本発明の一部に含まれ得る。特定の関連性には、たとえば、最大関連性、最小関連性、又は、最大関連性から最小関連性の範囲内の関連性が含まれる。この方法は、更に、選択された広告をユーザへ提示する。

【0010】本発明による他の実施例によれば、ユーザへ提示された一つ以上の文書の内容に基づいて、的確な広告をユーザへ送るシステムが提供される。このシステムは、一つ以上のユーザが選択可能な概念に対応した情報について文書の内容を解析するプロファイル内容理解器を含む。プロファイル内容理解器は、文書に含まれる概念の中で一つ以上のユーザが選択可能な概念に対応した概念を出力する。このシステムは、更に、一つ以上の広告主が選択可能な概念に対応した情報について文書の内容を解析する広告内容理解器を含む。比較器は、プロファイル内容理解器の出力と、広告内容理解器の出力とを比較し、複数の蓄積された広告の中から、上記比較に基づいて文書に含まれる情報に関連した広告を選択する。

【0011】本発明の更なる実施例によれば、一つ以上のユーザが選択可能な関心の概念に対応した情報の議論を識別するため文書を解析するコードをコンピュータに実行させ、ユーザへ提示された一つ以上の文書の内容に基づいて的確な広告をユーザへ送るコンピュータプログラムと、このコンピュータプログラムを記録したコンピュータ読み取り可能な記録媒体とが提供される。このコードは、ユーザによって定義された一つ以上の概念への文書の関連性の尺度を含むユーザ概念関連性指標を生成する。また、このコンピュータプログラムは、一つ以上の広告主が選択可能な関心の概念に対応した情報の議論を識別するため文書を解析するコードを含む。このコードは、広告主によって定義された一つ以上の概念への文書の関連性の尺度を含む広告主概念関連性指標を生成する。さらに、このコンピュータプログラムは、全体的な関連性を判定するため、文書についてのユーザ概念関連

性と広告主概念関連性を比較するコードを更に含む。また、このコンピュータプログラムは、特定の全体的な関連性を有する広告を選択するコードを含む。特定の関連性には、たとえば、最大関連性、最小関連性、或いは、最大関連性と最小関連性の範囲内の関連性が含まれる。

【0012】本発明の更に別の実施例によれば、広告対象を二つのウェブ対象に自動的に関連付ける方法が提供される。この方法は、少なくとも一つの第1の関心の概念に対応した情報を識別するため、一つ以上のウェブ対象を解析し、少なくとも一つの第2の関心の概念に対応した情報を識別するため一つ以上のウェブ対象を解析する。また、この方法は、第1の関心の概念と第2の関心の概念の間の関連性を判定する。さらに、この方法は、上記関連性に基づいて、一つ以上のウェブ対象を一つ以上の広告対象に関連付ける。

【0013】本発明の更に別の実施例によれば、広告対象を二つのウェブ対象に自動的に関連付けるコンピュータプログラムと、このコンピュータプログラムを記録したコンピュータ読み取り可能な記録媒体とが提供される。このコンピュータプログラムは、少なくとも一つの第1の関心の概念に対応した情報を識別するため一つ以上のウェブ対象を解析するコードと、少なくとも一つの第2の関心の概念に対応した情報を識別するため一つ以上のウェブ対象を解析するコードとを含む。このコンピュータプログラムは、第1の関心の関連と第2の関心の関連の間の関連性を判定するコードを含む。更にこのコンピュータプログラムは、上記関連性に基づいて、一つ以上のウェブ対象を、一つ以上の広告対象に関連付けるコードを含む。

【0014】本発明によれば、従来技術に対し多数の利点が得られる。本発明は、閲覧者の現在の関心に関連した広告をウェブページと共に一つの画面上に的確に出すことができる。本発明の実施例は、従来公知の技術よりも特定したターゲットを的確に決める。たとえば、一部の実施例は、あるユーザのスポーツの定義は、狩猟と釣りを含み、別のユーザのスポーツの概念はマラソン走を含むことを検出できる。このような実施例は、従来技術よりも巧く的確な広告を決めることができる。本発明の一部の実施例によれば、従来のウェブ・ブラウザ・ユーザ・インタフェースよりも優れた閲覧者のプライバシー保護が実現される。本発明の実施例によれば、従来技術よりも多くの情報がユーザへ提供される。本発明の多数の実施例によれば、最も関心があると考えられるリンクの指標を与えることにより、ユーザのウェブ閲覧体験の質が高められる。本発明の上述の利点及びその他の利点は、明細書の全体を通して明らかにされる。

【0015】本発明の特徴及び効果は、以下の実施例の説明及び添付図面を参照することによって、よりよく理解される。

【0016】本願は、1997年12月22日に出願された同一

出願人による係属中の米国特許出願第08/995,616号、発明の名称”AUTOMATIC ADAPTIVE DOCUMENT HELP SYSTEM”（代理人書類番号15358-004200US/ID-CRC-176）の全体を参考のため引用する。

【0017】また、本願と同日に出願された同一出願人による係属中の米国特許出願（代理人書類番号15358-005900US）、発明者Jamey Graham他、発明の名称”METHOD, SYSTEM AND COMPUTER CODE FOR CONTENT BASED WEB ADVERTISING”と、米国特許出願（代理人書類番号15358-005910US）、発明者Jamey Graham他、発明の名称”METHOD FOR INTRA-DOCUMENT COMPARISON IN A CONTENT BASED WEB ADVERTISING SYSTEM”を参考のため引用する。

【0018】

【発明の実施の形態】本発明によれば、閲覧者の関心に基づいてオンライン広告を閲覧者へ配信する関心プロフィールベースのターゲット販売システム、方法及びコンピュータプログラムが実現される。本発明の一実施例は、Reader's Helper™の名称で実施されている。

【0019】本発明の典型的な実施例は、たとえば、ウェブブラウザ上に一つ以上の販売情報領域を設ける。販売情報領域は、閲覧者が広告、公共サービス情報などにアクセスできるようにする。販売情報表示領域は、販売情報への一貫した控えめなアクセス方法を提供する。本発明による実施例は、たとえば、Phoenix, AZ.に本社が所在するNEOPLANT(www.neoplant.com)から供給されるウェブブラウザ、MICROSOFT CORPORATIONのInternet Explorer、NETSCAPE Navigator等のような当業者に公知の大多数の一般に普及したウェブブラウザを共に動作可能である。販売情報表示領域は、広告などのための場所も提供するが、この場所は文書の内容を邪魔しない。

【0020】的確に処理された広告とは、広告された製品又はサービスを購入する可能性が最も高い人々へ販売情報を組織的に提示することである。的確な処理は、たとえば、人の人口統計学的情報及びその他の情報に基づいて行われる。判定は、特定の人の製品若しくはサービスの使用状況に関する情報から明らかにされたその特定の人に対するプロフィールに基づいて行われる。たとえば、人が訪れたウェブページに関する情報は、その人の関心などのプロフィールを得るため、収集、解析される。特定のユーザを惹きつける可能性の高いウェブ広告対象は、ユーザのブラウザの販売情報領域に表示される。他の実施例の場合、販売情報は、人が使用するウェブチャネルに関する情報に基づいて配布される。一実施例において、特定のポータルを介して、ウェブページを訪問し、チャネルを使用するユーザは、ユーザのブラウザ内の販売情報領域を使用して広告を的確に処理するため利用可能な情報を与える。一部の実施例は、ユニフォーム・リソース・ロケータ(URL)を解析する。他の実施例は、ユーザの関心に関する情報を収集するため、ウェブページ自体の内容を解析する。

【0021】図1Aは、本発明の具体的な一実施例による販売情報を的確に処理する代表的なシステムの略構成図である。同図に示されたシステムは例示に過ぎず、請求項に記載された事項を制限するものではない。当業者は、この実施例を容易に変形、代替及び変更することができよう。図1Aには、1台以上のクライアントブラウザ12a、12b及び12cと、1社以上の広告主側の端末（以下、広告主と呼ぶ）14a、14b及び14cとの間にインタフェースを設けるサーバー10を含む。サーバー10は、各広告主のURLへアクセスする。URLは、特定の広告主の所望の販売情報を指定する。たとえば、図1Aにおいて、URL16a、16b及び16cは、それぞれ、広告主14a、14b及び14cの販売情報を指定する。現時点での好ましい実施例によれば、販売情報は、動画GIFフォーマットファイルで構成される。しかし、MPEG、JPEG等のフォーマットを有するマルチメディア対象、並びに、当業者に公知の他のフォーマットを有する情報対象が、本発明の一部の実施例で使用され得る。図1Aに示された典型的な実施例の場合に、データベース18は、たとえば、概念によって索引付けされた広告を蓄積する。図1Aは、広告主14a、14b及び14cからの販売情報を、それぞれ収容する複数の広告17a、17b及び17cを含む。広告17a、17b及び17cは、広告主が保守目的等のため広告にアクセスできるように対応したURL16a、16b及び16cを有する。サーバー10は、たとえば、プロフィールに収容されるユーザの関心と、広告主の関心と、閲覧されている現在の文書の内容との間の所定の関連性に基づいて、インターネット21を介して広告17a、17b及び17cを獲得することができる。

【0022】本発明による実施例の顕著な局面は、広告主がクライアントによる情報の閲覧によって実質的に妨げられることなく、対応した販売情報を更新できることである。好ましい一実施例によれば、広告主は、たとえば、使用状況を表す広告主の販売情報と対応した特定のウェブ対象へのヒットの回数をカウントすることによって、クライアントによる販売情報の使用状況を監視できる。一部の実施例の場合、使用状況統計量がサーバー10によって与えられる。

【0023】図1Bは、本発明の具体的な一実施例において、販売情報を的確に送る代表的な処理の簡略化されたフローである。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0024】図1Bには、たとえば、図1Aのサーバー10に組み込むことができる処理が示されている。文書100は、プロフィール内容理解器102と、広告内容理解器104の両方に入力される。プロフィール内容理解器102は、データベース19からユーザ選択可能概

念を入力する。プロファイル内容理解器102は、文書100の内容に関連したデータベース19からの一つ以上のユーザ内容概念を含む出力20を生成する。広告内容理解器104は、データベース18から広告主概念を入力する。広告内容理解器104は、文書100の内容に関連したデータベース18からの一つ以上の広告概念を含む出力22を生成する。概念比較器106は、プロファイル内容理解器102からユーザ概念出力20と、広告内容理解器104からの広告主概念22とを比較し、出力として、最良広告30を生成する。最良広告30は、文書の内容と、ユーザの関心と、広告の内容とに基づいて、クライアントブラウザ12aのユーザへの確に送られる広告である。ユーザ概念出力20は、ユーザの関心の概念に対する文書の関連性をユーザへ通知するため、ブラウザ12aへ供給され得る。好ましい一実施例において、これらの処理は、インターネットサービスプロバイダ(ISP)のサーバー、若しくは、プロキシサーバー上に同時に搭載され得るので、広告主は、データベース19内のユーザプロファイルへアクセスできない。具体的な一実施例では、広告主は、ISPから使用状況に関する統計的情報を取得することができる。

【0025】具体的な一実施例において、ユーザの個人的プロファイルは、ユーザ側に表示された文書の中で、個人的プロファイル内のユーザが選択した関心の概念によって定義されるようなユーザの関心が非常に高い特定の部分に注釈を付けるため使用される。一部の具体的な実施例の場合、閲覧されている文書の部分で注釈を付けられたフレーズ、及び／又は、閲覧されている文書の部分中の全てのフレーズは、ユーザが関心を持っている事項を正確に指摘するため使用される。たとえば、ユーザは、ユーザの関心の概念と関連した多数の強調されたキーワードを含む文書の一部を閲覧しているかもしれない。ユーザは、このような文書部分に非常に関心があるので、その内容を詳しく読む。したがって、広告は、ユーザが現在熟読中であるという情報に基づいて的確に送られる。この場合、強調された語は、適切な広告を同時に表示するため、広告概念と比較される。このような実施例は、ユーザが文書を読んでいるとき、ユーザにとってより好ましい広告を提供することができ、かつ、ユーザが文書中を移動するとき、表示される広告を変更することができる。

【0026】図1Cは、本発明の具体的な一実施例による典型的な検索技術の概略的な説明図である。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0027】図1Cには、ユーザが閲覧可能な例示的な文書100の第1部分105が示されている。文書の第1部分105は、幾つかの強調された語を含む。これらの強調された語は、第1の広告107(AD345)を表示さ

せる前後関係を構成する。その理由は、この広告の概念と、文書の第1部分105の概念との間に対応関係があるからである。その後、ユーザは文書の新しい部分である第2部分109を閲覧する。その中の強調された語は、第2の広告111(AD938)を表示させる前後関係を構成する。なぜならば、第2の広告111は、現在ユーザが閲覧中の概念に最も関連した広告である。強調された語だけを使用しても十分な関連性のあるスコアRjが得られない場合、閲覧中領域の他のフレーズが広告概念を収集するためのマッチング規準として使用される。

【0028】図2は、本発明の具体的な一実施例による典型的なシステムの構成図である。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。図2に示された例は、本発明と共に使用することができる多数の実現可能なコンピュータタイプ又はコンフィギュレーションの一例である。本発明による実施例は、ブラウザのような単一のアプリケーションプログラムで実現可能であるが、クライアント・サーバー関係のワークステーション、パーソナルコンピュータ若しくは遠隔端末のような分散コンピューティング環境における多数のプログラムとして実現してもよい。

【0029】図2は、ディスプレイ装置220と、ディスプレイ画面230と、キャビネット240と、キーボード250と、マウス270とを含むコンピュータシステム210である。マウス270及びキーボード250は、典型的なユーザ入力装置である。ユーザ入力装置のその他の例は、タッチスクリーン、ライトペン、トラックボール、データグローブ等である。図2は、本発明を実施するためのシステムの一例を表している。多数のシステムタイプ及びコンフィギュレーションが本発明と組み合わせ使用するため好適であることは、当業者に容易に認められる。

【0030】好ましい一実施例において、コンピュータ210は、Microsoft社のWindows^(R) NTオペレーティングシステムが動くPentium^(R) クラスのコンピュータである。しかし、本実施例の方法は、当業者によって、本発明の範囲を逸脱することなく、他のオペレーティングシステム及びアーキテクチャに容易に適合される。

【0031】マウス270は、ボタン280のような一つ以上のボタンを具備する。キャビネット240は、ディスクドライブ、プロセッサ、記憶手段などの一般的なコンピュータコンポーネントを収容する。本明細書中で使用される用語「記憶手段」は、ディスクドライブ、磁気テープ、固体メモリ、バブルメモリ等のようにコンピュータシステムと共に使用される任意の記憶装置を含む。キャビネット240は、コンピュータシステム210を、スキャナ260、外部記憶装置、他のコンピュータ或いは付加的な周辺機器のような外部装置に接続する入出力(I/O)インタフェースカードのような付加的なハ

ードウェアを収容する。

【0032】図3は、図2のコンピュータシステムの基本サブシステムの構成図である。図3において、サブシステムは、中央プロセッサ300、システムメモリ310などのブロック形式で示されている。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0033】サブシステムは、サブシステムバス320によって相互連結される。同図には、プリンタ、キーボード、固定ディスク、及び、その他のブロックのような付加的なサブシステムが示されている。周辺装置及び入出力(I/O)装置は、シリアルポート330のような従来技術において公知の多数の手段を用いてコンピュータシステムに接続される。たとえば、シリアルポート330は、コンピュータシステムをモデム、マウス入力装置、又は、スキャナに接続するため使用される。システムバス320による相互連結は、中央プロセッサ330が各サブシステムと通信し、システムメモリ310若しくは固定ディスク332からの命令の実行を制御することを可能にし、並びに、サブシステム間での情報の交換を可能にする。この他のサブシステムの配置及び相互連結は、当業者であれば容易に実現することができる。システムメモリ310及び固定ディスク332は、コンピュータプログラムを記憶するための具体的な媒体の例であり、具体的な媒体の例には、フロッピーディスクと、着脱可能式ハードディスクと、CD-ROM及びバーコードのような光学的記憶装置と、フラッシュメモリ、読み出し専用メモリ(ROM)及びバッテリーバックアップメモリのような半導体メモリとが含まれる。

【0034】好ましい一実施例によれば、一つ以上の概念に対する文書の内容の関連性の指標を与えるため概念指標器が、文書及び販売情報に一つ以上の表示領域を準備し得るウェブブラウザと共に使用される。概念指標器は、ユーザ選択可能な関心の概念に対するユーザが閲覧した文書の内容の関連性についての解析結果を示す。値は、ユーザがその文書に対し関心があるかどうかについて判定された尤度に基づいて割り当てられる。好ましい一実施例において、文書全体を通して使用されたキーワードが、ユーザ選択可能な関心の概念に対する関連性に応じて自動的に強調されるので、ユーザは非常に効率的に関連性情報を見つけることができる。ユーザが定義した各関心の概念は、その概念に関する文書内の重要な事項を記述するキーワードの集まりを含む。概念指標器は、文書がその概念に関連しているか、並びに、読むだけの価値があるかどうかについての指標をユーザに与えることができる。概念指標器については、参考のため引用された、本願と同一出願人による同時係属中の米国特許出願第08/995,616号に記載されている。

【0035】本発明による実施例は、製品に関心を持つ

ている人によって読まれる可能性の高いウェブページ又は文書の閲覧者への確に向けられる広告並びにその他の販売情報を作成する能力を提供する。たとえば、本発明による技術は、雑誌Outsideの広告が、登山若しくはトレッキングに関する情報を含む文書又はウェブページの読者、並びに、それらの概念に好みを示す読者の興味を惹くという判定をすることができる。たとえば、会議についての広告は、同一の主題に関する前回の会議からの論文の要約を表示することによって興味を惹くことができる。

【0036】本発明の特定の典型的な実施例によれば、広告は、ユーザの関心の概念と、これらの概念が特定の文書と関連する度合いとに基づいて的確に向けられる。関連性は、後述のスコアリング若しくはその他の方法によって決定される。この情報は、実際の文書の内容と組み合わせられて、広告を的確にユーザへ向けるため使用される。関心の組を定義する概念は、ユーザ入力又はその他のメカニズムを用いて収集される。好ましい一実施例によれば、ユーザのプライバシーは、ユーザのプロファイルが広告主と共有されないために保護される。ユーザプロファイルは、たとえば、クライアントコンピュータ、ISPサーバー、又は、プロキシサーバーに局所的に蓄積される。一部の具体的な実施例の場合、広告主には、たとえば、ユーザのブラウザによって記録された統計的情報が提供される。

【0037】本発明の更に別の実施例によれば、広告は内容解析に基づいて選択される。広告は、概念が割り当てられたシステムに基づくポータルなどに対し、ウェブページ内容と関連付けられる。たとえば、ウェブポータルのプロバイダは、ポータルウェブページにリンクされたウェブページと広告との間の関連性の組を作成することができる。特定の一実施例によれば、内容解析技術は、特定の 카테고리に基づいて広告対象とウェブ対象の間の関連性を与える。

【0038】図4Aは、本発明の特定の代表的な実施例において、ユーザへ提示された一つ以上の文書に基づいて、販売情報をユーザへの確に向けられる簡単化された処理のフローチャート401である。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0039】図4Aのステップ402において、一つ以上のユーザ選択可能な関心の概念に対応した議論を識別するため、文書を解析する。文書の内容は、概念と文書の間の類似性を表す概念毎に確率を生成するため、文書のテキストが概念の集まりと比較される図5乃至8を参照して後述する技術を用いて解析され得る。関連性判定技術に関する更なる説明は、参考のため引用した本願と同一出願人による米国特許出願第08/995,616号に記載されている。

【0040】次に、ステップ404において、関連した概念の集まりはステップ402で判定された確率に基づいて生成される。

【0041】ステップ406において、文書の内容は、文書と、広告主の定義した概念との関連性を判定するため、ステップ402と同じ技術を用いて再度解析される。特定の一実施例では、これらの広告概念は、広告主がたとえば自分の製品と関連性があるとして確認したキーワードである。特定の一実施例において、確率推定値が文書毎に各広告概念について生成される。

【0042】次に、ステップ408において、関連性概念の集まりは、ステップ406で判定された確率推定値に基づいて生成される。

【0043】選択的なステップ409において、ステップ404で生成された概念の集まりが、文書の関連性の標識をユーザへ表示するため構成されたウェブブラウザへ配布される。

【0044】次に、ステップ410において、ステップ404で与えられたユーザ概念の集まりと、ステップ408で与えられた広告概念の集まりとの間で、たとえば、いずれの広告概念がユーザ概念に最も類似しているかを判定するため、比較が行われる。

【0045】ステップ412では、選択された広告がユーザへ配布される。選択的に、広告は、たとえば、ウェブブラウザの販売情報領域に表示される。

【0046】図4Bは、本発明の特定の代表的な実施例にしたがって、広告対象を、ウェブ対象及び図4Aのステップ410におけるユーザ概念と自動的に関連付ける単純化された処理のフローチャートである。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0047】図4Bのステップ422は、図4Aのステップ404で獲得されたユーザ概念と、図4Aのステップ408で獲得された関連した広告概念の集まりとの間の類似性解析を実行する。

【0048】一部の具体的な実施例の場合、図4Bのステップ422で行われるような比較を実行するため、当業者に周知の技術が使用される。たとえば、文献：Robert R. Korfhage: "Information Storage and Retrieval", John Wiley & Sons, Inc., 1997には、マッチングのため使用される幾つかの関連性測定量が示されている。

典型的な一実施例によれば、マッチング技術は、各ユーザ及び広告概念が文書、すなわち、用語の集まりを表現する文書類似性方法を使用する。一部の具体的な実施例では、ユーザ概念及び広告概念の両方に対するベクトル表現が作成される用語—頻度測定が使用される。これらのベクトルは、たとえば、1が一致を示し、0が不一致を示すターゲットベクトルのサイズを持つ新しいベクトルを作成するため比較される。典型的な一実施例において、ユーザ概念Cは、用語のベクトル

$$C = \langle t_1, t_2, t_3, \dots, t_x \rangle$$

として表現され、広告概念Aは、用語のベクトル

$$A = \langle t_1, t_2, t_3, \dots, t_y \rangle$$

として表現される。ベクトルA_jにおけるベクトルC_jからの用語の頻度を示すベクトル比較によって、ベクトル

$$V_{ij} = \langle 0, 1, 1, 0, 1, 0, 0 \rangle$$

が得られる。

【0049】本例の場合、概念C_iは、7個のキーワードを含み、その中の3個が広告概念A_jと一致している。ここで、W_{ij}がベクトルV_{ij}内の値の合計を表すとする。この比較の類似性値は、S_{ij}によって表現され、W_{ij}をV_{ij}の長さ(Xで表す)で割った値に、文書とユーザ概念との間で図4Aのステップ402で生成された比較値c_i=0.56を乗算した値に一致し、

【0050】

【数5】

$$S_{ij} = \left(\frac{W_{ij}}{X_i} \right) \times c_i$$

であり、たとえば、

$$(3/7) \times 0.56 = 0.24$$

となる。

【0051】値c_iは、より高いスコアをユーザ概念につける方向で、S_{ij}にバイアスを与える。全ての関連したユーザ概念及び全ての関連した広告概念が与えられた場合、ユーザ概念／広告概念の組み合わせ毎に類似性値を有するマトリックスを構築することができる。表1は、具体的な一実施例におけるこのようなマトリックスの典型的な例である。

【0052】

【表1】

表1：比較マトリックス

概念A _d	C ₁	C ₂	C ₃	...	C _i	関連性
A ₁	S _{1,1}	S _{2,1}	S _{3,1}	...	S _{m,1}	R ₁
A ₂	S _{1,2}	S _{2,2}	S _{3,2}	...	S _{m,2}	R ₂
A ₃	S _{1,3}	S _{2,3}	S _{3,3}	...	S _{m,3}	R ₃
...
A _j	S _{1,n}	S _{2,n}	S _{3,n}	...	S _{m,n}	R _n

この典型的な例の場合に、関連したユーザ概念の集合 $\{C_1, C_2, C_3, \dots, C_i\}$ と、関連した広告概念の集合 $\{A_1, A_2, A_3, \dots, A_j\}$ とが含まれる。個別の類似性値は、ユーザ概念と広告概念の間の比較毎に S_{ij} として表される。ここで、 i はユーザ概念添え字であり、 j は広告概念添え字である。

【0053】次に、ステップ424において、与えられた全てのユーザ概念に対する最良の広告概念が判定される。このステップでは、図4Aのステップ406で広告と文書の間で生成された比較値 a_j は、広告概念が現在の文書により類似する方向で、ユーザ／広告概念比較値 S_{ij} にバイアスを加える。各広告概念に対するこの組み合わせは、広告を格付けする最終的な類似度測定量を生ずる。

【0054】

【数6】

$$R_j = \left(\sum_i S_{ij} \right) \times a_j$$

同じ値 R_j を有する多数の広告が存在する場合、値 a_j は引き分けを避けるために使用され、文書に関してより高いスコアを有する広告概念がユーザのブラウザに配布される。 R_j がある閾値、たとえば、20%の関連性を超えないとき、広告は、元の内容ベースの値 a_j を用いて格付けされる。この場合、より良い解決法は、内容に関して現在の文書に少なくとも類似している文書を配布することである。値 R_j も値 a_j もどちらも閾値を越えない場合、広告は表示されず、或いは、関連したヘッドラインニュース、又は、ユーザのプロファイルに関連したニュースのような他の情報が広告領域に表示される。

【0055】図5は、本発明の具体的な実施例において、電子的に蓄積された文書に関して解析する処理の簡単化されたフローチャートである。この図に示された処理は実施例であり、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。ウェブ対象等の電子的に蓄積された文書を解析する方法は、参考のため引用した本願と同一出願人による同時係属中の米国特許出願第08/995,616号に非常に詳しく記載されている。

【0056】文書100は、電子フォーマットで蓄積される。文書100は、最初にスキャン処理される場合がある。或いは、文書は、たとえば、HTML、Postscript、LaTeX、その他のワードプロセッシングフォーマット、若しくは、電子メールフォーマットでもよい。以下に説明する実施例では、HTMLフォーマットが使用されるが、別の具体的な実施例では、本発明の範囲を逸脱することなく、HTML以外のフォーマットを使用することができる。ユーザ504は、ブラウザ506を使って文書100へアクセスする。ブラウザ506は、好ましくは、NE TSCAPE Navigator又はMICROSOFT CORPORATION Internet

Explorerのようなハイパーテキストブラウジングプログラムであるが、たとえば、通常のワードプロセッシングプログラムでもよい。

【0057】付注エージェント508は、ブラウザ506による閲覧に備えるため、文書100に注釈を付ける。付注エージェント508による処理は、テキスト処理段510と、内容理解段512と、フォーマッティング段514の3段で構成される。テキスト処理段510への入力、たとえば、未加工テキストでも、又は、htmlタグが埋め込まれたテキストでもよい。テキスト処理段510からの出力、及び、内容理解段512への入力、解析済みテキストストリームである。この解析済みテキストストリームは、除去された特定の語若しくは句（フレーズ）の周りの特殊タグのようなフォーマッティング情報を含むテキストストリームである。内容理解段512からの出力、及び、フォーマッティング段514への入力、注釈付きテキストストリームである。フォーマッティング段514の出力は、印刷若しくはブラウザ506による閲覧が可能なフォーマット済みテキストファイルである。

【0058】付注エージェント508の処理は、好ましくは、ランタイム処理である。付注は、好ましくは、予めテキストに挿入されるのではなく、ユーザ504が文書100を閲覧する要求をしたときに生成される。したがって、付注は、好ましくは、動的処理である。他の実施例の場合、付注エージェント508は、バッチプロセスとして、バックグラウンドで動作する。

【0059】付注エージェント508によって付加される注釈は、ユーザ504により選択された関心の概念に依存する。ユーザ504は、文書100内での関心の概念の議論の場所を識別するため、付注エージェント508によって使用される情報を入力する。好ましい一実施例において、この情報は、バイズ信頼ネットワークの構造を定義する。関心の概念及びその他のユーザ専用情報は、ユーザプロファイルファイル516に保持される。ユーザ504は、ユーザプロファイルファイル516の内容を変更するため、プロファイルエディタ518を利用する。

【0060】図6Aは、詳細に示されたテキスト処理段510を含む図5の自動注釈付けソフトウエアのアーキテクチャを表す図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0061】図6Aには、文書100のソースがネットワーク602を介してアクセスされることが示されている。実現可能なソースには、たとえば、インターネット604と、イントラネット606と、文書画像を捕捉するデジタルコピー機608と、ファクシミリ機、スキャナ、プリンタ等のその他の事務機器610とが含まれ

る。その他のソースは、ユーザ自身のハードドライブ332である。

【0062】テキスト処理段510は、ファイル入出力段612と、更新段614と、言語処理段616とを含む。ファイル入出力段は、ネットワーク602から文書を読む。更新段614は、履歴ファイル618に、最近閲覧された文書の履歴を保持する。言語処理段616は、文書100のテキストを解析し、テキスト処理段510の解析されたテキスト出力を生成する。

【0063】図6Bは、内容理解段512が詳細に示されている図5の自動注釈付け処理のフローチャートである。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0064】パターン識別段620は、テキスト処理段510の解析されたテキスト出力中の特定のパターンを検索する。検索される特定のパターンは、ユーザファイル516の内容によって判定される。パターンが見つけた後、注釈タグが注釈タグ付加段622によって、パターンの場所を示すため、解析済みテキストに付加される。好ましいHTML形式の実施例の場合、これらの注釈タグはHTMLフォーマットと互換性がある。しかし、タグ付け処理は、LaTeX、Postscript等にも適用される。プロファイル更新段624は、注釈タグ付加段622の出力を監視し、関心の概念の場所の周辺にあるテキストを解析する。図7を参照して更に説明されるように、ユーザプロファイル516の内容の変更は、この周辺テキストの解析に基づいている。その効果は、注釈付けの性能を高めるため、パターン識別段620によって検索されるパターンを自動的に更新することである。

【0065】図6Cは、本発明の具体的な一実施例における図5の自動注釈付け処理のフローチャートである。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0066】フォーマッティング段514は、ブラウザ506による閲覧が容易に行えるように、内容理解段512によって供給された注釈付きテキストをフォーマッティングするテキスト翻訳段626を含む。パターン識別段620は、関心のキーワード及びキーフレーズを検索し、見つけられたキーワードに基づいて関連した概念の議論の場所を見つける。キーワードの識別と、関連した議論を見つけるためのキーワードの利用は、好ましくは、推論システムに問い合わせることによって実現される。推論システムは、好ましくは、ベイズ信頼ネットワークである。

【0067】図7には、本発明の具体的な一実施例におけるパターン識別段620で使用されるような推論システムを実現する典型的なベイズ信頼ネットワーク700の一部分が示されている。この図に示された例は一実施

例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0068】同図において、具体的なユーザが指定した関心の概念、すなわち、第1の概念は、楕円702で示されている。第1の概念702によって指定された概念に関連した二次概念は、楕円704で示されている。第1の概念702と、各二次概念704とを連結する線は、二次概念の議論が第1の概念の議論を意味することを示している。一つの二次概念704と第1の概念702のコネクションには、パーセント表示された確率値が関連付けられる。二次概念の議論は、一つ以上のキーワード若しくはキーフレーズ（図7に示されない）によって指定される。

【0069】ベイズ信頼ネットワーク700の構造は、本発明に適用可能な唯一の実現可能な構造ではない。たとえば、二次概念の存在が三次概念の存在によって示され、以下同様に続く3レベル以上の階層を有するベイズ信頼ネットワークを利用してもよい。上記の好ましい実施例の場合、キーワード若しくはキーフレーズの存在は、常に、二次概念の議論の存在を示すが、キーワード若しくはキーフレーズの存在が確率付きの二次概念の議論を示すように信頼ネットワークを構成することが可能である。

【0070】概念と、キーワード及びキーフレーズと、相互連結と、確率とを含むベイズ信頼ネットワーク700の構造に対する1次ソースは、ユーザプロファイルファイル516である。好ましい一実施例において、ユーザプロファイルファイル516は、多数のユーザのプロファイルの中から、編集及び使用のため選択可能である。

【0071】推論システム700の構造は、注釈付けシステムの使用中に変更可能である。変更は、バックグラウンドで自動的に行われ、或いは、明示的なユーザフィードバック入力に関与する。パターン識別段620によって判定された関心の概念の場所は、プロファイル更新段624によって監視される。プロファイル更新段624は、監視の概念の場所に関して解析された文書内で他のキーワード及びキーフレーズの付近に注意する。特定のキーワード及びキーフレーズが常に関心の概念の近傍にあるとき、推論システム700の構造及び内容は、ユーザ入力を用いることなくバックグラウンドで、プロファイル更新段624によって更新される。これは、確率値を変更し、二次概念と概念の間に新しいコネクションを導入し、或いは、新しいキーワードを導入することを意味する。本明細書中で、キーワードは単一のキーワードにより構成され、句は、一つ以上の語、すなわち、キーフレーズにより構成される。

【0072】ユーザ504は、語（ワード）又は句（フレーズ）がキーワード又はキーフレーズとして定義され

ていなくても、特定の概念に関連している文書100中の語又は句を選択する。推論システム700は、新しいキーワードを組み込むよう更新される。

【0073】ユーザ504は、既存のキーワードに関するフィードバックを与え、関心の概念に対する認識されたキーワードの関連性を示す。選択されたキーワードが、関心の概念に非常に関連していることが示された場合、選択されたキーワードによって指定される二次概念を関心の概念に結び付ける確率値は増加する。一方、ユーザ504が、選択されたキーワードに余り関心が無いことを示した場合、コレラのキーワードを概念に結び付ける確率値は減少する。

【0074】図8は、本発明の具体的な実施例においてユーザプロフィールを定義するユーザインタフェースを示す図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0075】ユーザインタフェース画面800は、プロフィールエディタ518によって与えられる。プロフィール名ボックス802は、ユーザが編集されるプロフィールを割り当てる個人名若しくはグループ名を入力するため利用される。これにより、本発明による注釈付けシステムは、特定のユーザ若しくはグループへ個人化される。パスワードボックス804は、プロフィール編集操作の前に正しいパスワードの入力を要求することによって、安全性を確保する。

【0076】定義済み概念リスト806は、ユーザプロフィールに既存の全ての概念を一覧する。概念追加ボタン808を選択することにより、ユーザは新しい概念を追加する。概念編集ボタン810を選択することにより、ユーザは、現在選択されている一覧された概念に適するように信頼ネットワークを変更する。削除ボタンを選択することにより、ユーザは概念を削除してもよい。

【0077】概念が編集のため選択されたとき、その名前は概念名ボックス813に現れる。選択された概念に属する信頼ネットワークの一部分は、信頼ネットワーク表示ウィンドウ814に示される。信頼ネットワーク表示ウィンドウ814は、選択された概念、選択された概念に関連しているとして定義された二次概念、及び、関係ごとに関連付けられたパーセンテージ値とを表示する。ユーザは、二次概念追加ボタン815を選択することにより、二次概念を追加してもよい。ユーザは、信頼ネットワーク表示ウィンドウ814内の二次概念を選択することにより、二次概念を編集する。二次概念削除ボタン818は、ユーザが信頼ネットワークから二次概念を削除できるようにする。

【0078】二次概念追加ボタン815を選択することにより、二次概念追加ウィンドウ820が現れる。二次概念追加ウィンドウ824は、新しい二次概念名を入力

する二次概念名ボックス822を含む。スライドコントロール824は、新たに選択された二次概念が現れる場合に、ユーザが選択された概念の出現確率を定義するパーセンテージ値を選択できるようにする。キーワードリスト826は、二次概念の議論を示すキーワード及びキーフレーズを一覧する。ユーザは、新しいキーワード若しくはキーフレーズを入力するための(図示されない)ダイアログボックスを表示させるキーワード追加ボタン828を選択することにより、このリストに追加する。ユーザは、キーワード若しくはキーフレーズを選択し、次に、キーワード削除ボタン830を選択することにより、このキーワード若しくはキーフレーズを削除する。ユーザが新しい二次概念を定義し終えたとき、OKボタン832を選択することにより、その定義を確認する。二次概念編集ボタン816の選択によって、二次概念追加ウィンドウ820に類似したウィンドウが表示され、選択された二次概念の再定義が可能になる。

【0079】バックグラウンド学習検査ボックス836が選択されているかどうかを判定することにより、ユーザはプロフィール更新段624の動作を許可若しくは禁止する。ウェブ自動フェッチ検査ボックス838は、ユーザが自動ウェブ探索処理を許可するかどうかを選択できるようにする。このウェブ探索処理が許可されたとき、特定のキーワード若しくはキーフレーズが定義された概念が議論されていると判定された場所の近傍で頻繁に検出されるならば、ALTAVISTAのようなウェブ検索ツールが、このキーワード若しくはキーフレーズを含む文書をワールド・ワイド・ウェブ上で検索するため利用される。閾値スライドコントロール840は、ユーザがこの自動フェッチ処理のための閾値関連レベルを設定できるようにするため設けられる。

【0080】図9A及び9Bは、本発明の具体的な実施例においてフィードバックを行うユーザプロフィールを定義するユーザインタフェースを示す図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0081】ユーザは、任意のテキストを選択し、第1のフィードバックウィンドウ902を呼び出す。テキストは、注釈付けシステムによって関連していることが予め識別されていても識別されていなくてもよい。図9Aに示された第1のフィードバックウィンドウ902において、ユーザ504は、選択されたテキストが関連している概念を示す。第1のフィードバックウィンドウ902は、既に信頼ネットワーク700の一部であるキーワード若しくはキーフレーズに対する関連性レベルを調節するときには必要とされない。ユーザが第1のフィードバックウィンドウ902内で概念を選択したとき、第2のフィードバックウィンドウ904が関連性の程度を選択するため表示される。図9Bにおける第2のフィード

バックウィンドウ904は、関連性のレベルに対する3種類の選択肢、良、並（不確か）、悪を与える。或いは、スライドコントロールは、関連性のレベルを設定するため使用される。選択されたテキストが未だ信頼ネットワーク700内のキーワード若しくはキーフレーズではない場合、新しい二次概念が関連したキーワードと共に追加される。選択されたテキストが既にキーワードであるとき、信頼ネットワーク700内の確率値は、このユーザフィードバックに応じて適切に変更される。

【0082】図10は、本発明の一実施例において処理されるHTML文書100の一部分を示す図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易に成し得るであろう。

【0083】関連したテキストを含むテキストのブロックの前には、図10の…タグ1006aのような…タグと、図10の…タグ1008aのようなタグとが先行する。…タグは、注釈付けされた文書が関連した概念の名前と、選択的に二次概念の名前とを含む。たとえば、…タグ1006aは、…タグ1006aと対応したタグ1008aの間のテキストが概念"Agent"に関連していることを示すテキスト"class=rhtopic_Agent"を含む。テキスト"intelligent agent"は、典型的な本例における特定の概念に関連したテキストを含む。

【0084】好ましい一実施例において、本発明を実施するソフトウェアは、Java言語で記述される。好ましくは、ソフトウェアは、スタンドアロン型ブラウザプログラムの一部を形成する。或いは、そのコードは、上記の注釈タグを含むHTML文書を閲覧するため使用されるウェブブラウザと共に動作するプラグイン形式でもよい。

【0085】

【実施例】図11Aは、本発明の具体的な実施例におけるウェブブラウザの典型的な販売情報表示領域の例の説明図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易に成し得るであろう。

【0086】図11Aは、二つの広告表示領域を有するブラウザ画面1503を示す。ハンドヘルド型分散コンピューティングに関する会議用の第1の広告1512と、ACMデジタルライブラリに関する第2の広告1514が、ウェブブラウザ画面1503によって表示される。この具体的な例の場合に、両方の広告表示領域1512及び1514は、関連した概念、ウェアラブルコンピュータに関する広告を表示する。領域1514に表示された広告は、領域1512に表示された広告よりもウェアラブルコンピュータの基礎的な概念の広い局面に向けられることに注意すべきである。これらの広告は、上述の関連性判定技術を用いて選択される。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置

が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。

【0087】図11Bは、本発明の具体的な実施例におけるウェブブラウザの別の典型的な販売情報表示領域の例の説明図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易に成し得るであろう。

【0088】図11Bは、ユーザインタフェースツールに関する技術文書が表示されている表示領域1505を示す。販売情報表示領域1522及び販売情報表示領域1524に表示された広告は、ユーザインタフェース上で二つの異なる会議を指示する。これらの広告は、関連性判定技術を用いて選択される。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。

【0089】図11Cは、本発明の具体的な実施例におけるウェブブラウザの更に別の典型的な販売情報表示領域の例の説明図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易に成し得るであろう。

【0090】図11Cは、NCAAバレーボールに関する記事が掲載されたブラウザ画面1507を示す。バレーボール器具供給元及びESPNに関する広告は、それぞれ、第1の表示領域1532及び第2の表示領域1534に表示される。広告は、関連性判定技術を用いて選択される。ユーザインタフェースツールに関する技術文書が表示されている表示領域1505を示す。販売情報表示領域1522及び販売情報表示領域1524に表示された広告は、ユーザインタフェース上で二つの異なる会議を指示する。これらの広告は、関連性判定技術を用いて選択される。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。される。

【0091】図11Dは、本発明の具体的な実施例におけるウェブブラウザの典型的な販売情報表示領域の例の説明図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易に成し得るであろう。

【0092】図11Dは、たとえば、NEOPLANETにより

提供されるブラウザからの画面であるブラウザ画面 1509 を示す。当業者は、MICROSOFT CORPORATION による Internet Explorer、NETSCAPE Navigator のような種々のブラウザを使用することによって多数の実施例を作成することができる。ブラウザ画面 1509 は、表示領域に技術文書を掲載している。SUN MICROSYSTEMS の広告がブラウザ画面 1509 の広告領域 1544 に表示される。この広告は、関連性判定技術を用いて選択される。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。

【0093】図 11E は、本発明の具体的な実施例におけるウェブブラウザの更に別の典型的な販売情報表示領域の例の説明図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0094】図 11E は、ツールバー 1552 を有するブラウザ画面 1511 を示す。ブラウザ画面 1511 は、MICROSOFT CORPORATION から供給される Internet Explorer 5.0 であり、ツールバー 1552 は、YAHOO, Incorporated によるツールバーである。ツールバー 1552 は、広告用の販売情報表示領域 1554 を有する。表示画面 1511 は、科学推理小説及び映画、“2001 Space Odyssey”に関する文書を表示する。販売情報表示領域 1554 は、映画“Star Wars”の広告を掲載する。この広告は、関連性判定技術を用いて選択される。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。

【0095】図 11F 及び図 11G は、本発明の具体的な実施例による概念選択及び広告的確化処理の動作を説明する図である。この図に示された例は一実施例に過ぎず、本発明の範囲を制限するものではない。当業者は、本実施例の変形、代替及び変更を容易になし得るであろう。

【0096】図 11F は、複数の概念 1502 を有するブラウザ画面 1513 を示す。各概念は、ユーザがその概念に興味をもっていることを示すため、ユーザによって選択されている。ブラウザ画面 1513 は、選択された概念に関連した二つの広告を表示する広告表示領域 1504 を含む。第 1 の広告は、HCI 国際会議のための広告である。第 2 の広告は、UIST '99 会議のための広告である。対照のため、図 11G は、同様のブラウザ画面 1513 を示す。しかし、図 11G では、ユーザは、複数の概念 1502 の中で、Interface と AugReal (Augmented Reality) の二つの概念を抑止している。選択された関心の概念の変更に応じて、広告表示領域 1504 は、二

つの異なる広告を表示する。第 1 の広告は、分散型ハンドヘルドコンピューティング会議のための広告である。第 2 の広告は、IBM に関する広告である。ユーザは関心の概念の選択を変更しているので、表示された二つの広告は、現時点で、ユーザへの関連性が最も高い。これらの広告は、関連性判定技術を用いて選択されている。しかし、表示領域に異なる概念を有する実施例、或いは、表示領域の配置が異なり、表示領域の数が異なるなどの別の実施例は、本発明の範囲を逸脱することなく当業者によって容易に作成される。

【0097】最後に、本発明は、広告を的確に送るための文書内容解析方法を提供する。上記の説明において、本発明は、具体的な実施例に基づいて説明されている。これらの実施例の変形又は変更は容易に行える。たとえば、表示中の文書画像のサイズ及び配置の変更、文書画像の見え方及び形状の変更、文書画像等の操作中における聴覚的効果の付加、文書画像の内容を説明するオーディオメモの追加、他の推論システムによるベイズ信頼ネットワークの置換などは、本発明の他の実施例に組み込まれる。しかし、実施例の多数の置換及び変更は、特許請求の範囲に記載された事項、並びに、その均等物の広い精神及び範囲に含まれる。

【0098】明らかに、本発明の多数の置換及び変形が上記の教示に関して実現可能である。したがって、請求項に記載された事項の範囲内で、本発明は具体的に開示された形態に限定されることなく実施されることに注意する必要がある。

【図面の簡単な説明】

【図 1A】本発明の具体的な一実施例を実現する代表的なシステムの構成図である。

【図 1B】本発明の具体的な一実施例を実現する代表的なシステムの構成図である。

【図 1C】本発明の具体的な一実施例を実現する代表的なシステムの構成図である。

【図 2】本発明の具体的な一実施例におけるコンピュータシステムの概要図である。

【図 3】図 2 に示されたコンピュータシステムの基本サブシステムのブロック図である。

【図 4A】本発明の具体的な一実施例における処理ステップの略フローチャートである。

【図 4B】本発明の具体的な一実施例における処理ステップの略フローチャートである。

【図 5】本発明の具体的な一実施例における自動注釈付け処理の略ブロック図である。

【図 6A】本発明の具体的な一実施例における自動注釈付け処理の略ブロック図である。

【図 6B】本発明の具体的な一実施例における自動注釈付け処理の略ブロック図である。

【図 6C】本発明の具体的な一実施例における自動注釈付け処理の略ブロック図である。

【図 7】本発明の具体的な一実施例における自動付注で利用される代表的なベイズ信頼ネットワークの概略図である。

【図 8】本発明の具体的な一実施例においてユーザプロフィールを定義するための代表的なユーザ・インタフェースの説明図である。

【図 9 A】本発明の具体的な一実施例においてユーザ・フィードバックを行う代表的なインタフェースの説明図である。

【図 9 B】本発明の具体的な一実施例においてユーザ・フィードバックを行う代表的なインタフェースの説明図である。

【図 1 0】本発明の具体的な一実施例において処理される代表的な HTML 文書の一部の説明図である。

【図 1 1 A】本発明の具体的な一実施例における代表的な画面表示の説明図である。

【図 1 1 B】本発明の具体的な一実施例における代表的な画面表示の説明図である。

【図 1 1 C】本発明の具体的な一実施例における代表的な画面表示の説明図である。

【図 1 1 D】本発明の具体的な一実施例における代表的な画面表示の説明図である。

【図 1 1 E】本発明の具体的な一実施例における代表的な画面表示の説明図である。

【図 1 1 F】本発明の具体的な一実施例による概念選択及び広告的確化処理の動作を説明する図である。

【図 1 1 G】本発明の具体的な一実施例による概念選択及び広告的確化処理の動作を説明する図である。

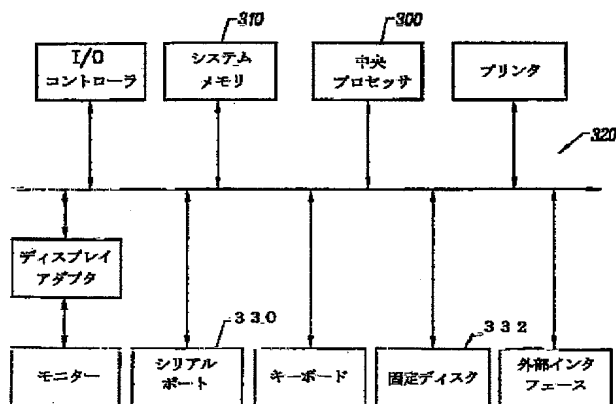
【符号の説明】

1 8 広告主概念データベース

1 9 ユーザ概念データベース
2 0 ユーザ概念
2 2 広告主概念
3 0 広告
1 0 0 文書
1 0 2 プロファイル内容理解器
1 0 4 広告内容理解器
1 0 6 概念比較器
5 0 4 ユーザ
5 0 6 ブラウザ
5 0 8 付注エージェント
5 1 0 テキスト処理段
5 1 2 内容理解段
5 1 4 フォーマット段
5 1 6 ユーザプロフィール
5 1 8 プロファイルエディタ
6 0 2 ネットワーク
6 0 4 インターネット
6 0 6 イン트라ネット
6 0 8 コピー機
6 1 0 その他の機器
6 1 2 ファイル入出力
6 1 4 更新段
6 1 6 言語処理段
6 1 8 履歴
6 2 0 パターン識別段
6 2 2 注釈タグ付加段
6 2 4 プロファイル更新段
6 2 6 テキスト翻訳段
7 0 0 ベイズ信頼ネットワーク

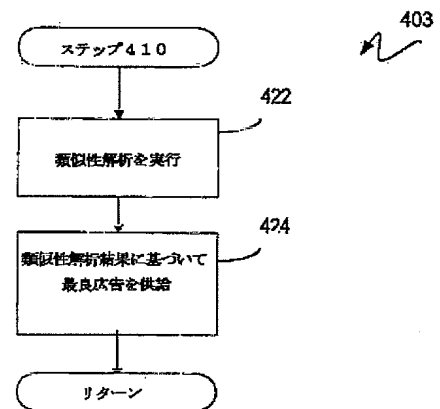
【図 3】

図 2 に示されたコンピュータシステムの基本サブシステムのブロック図



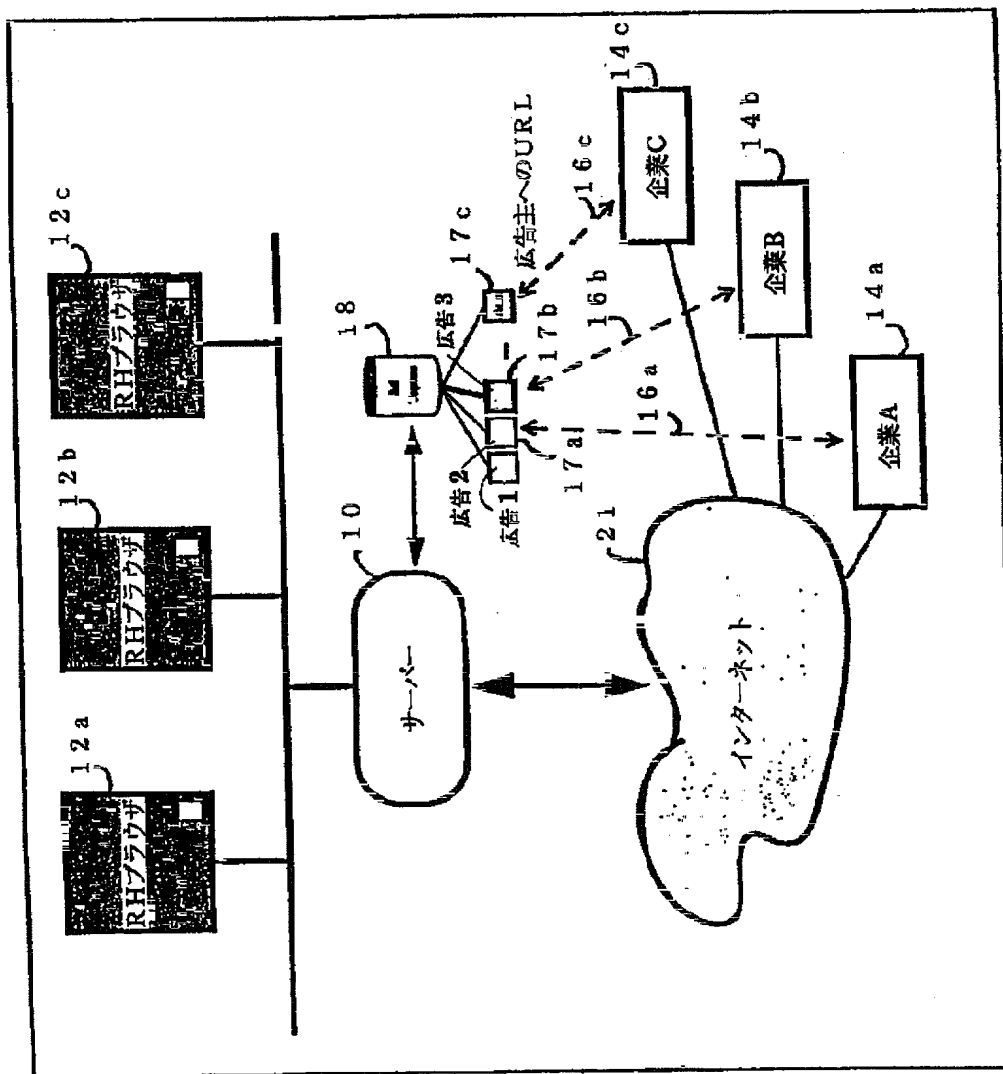
【図 4 B】

本発明の具体的な一実施例における処理ステップのフローチャート



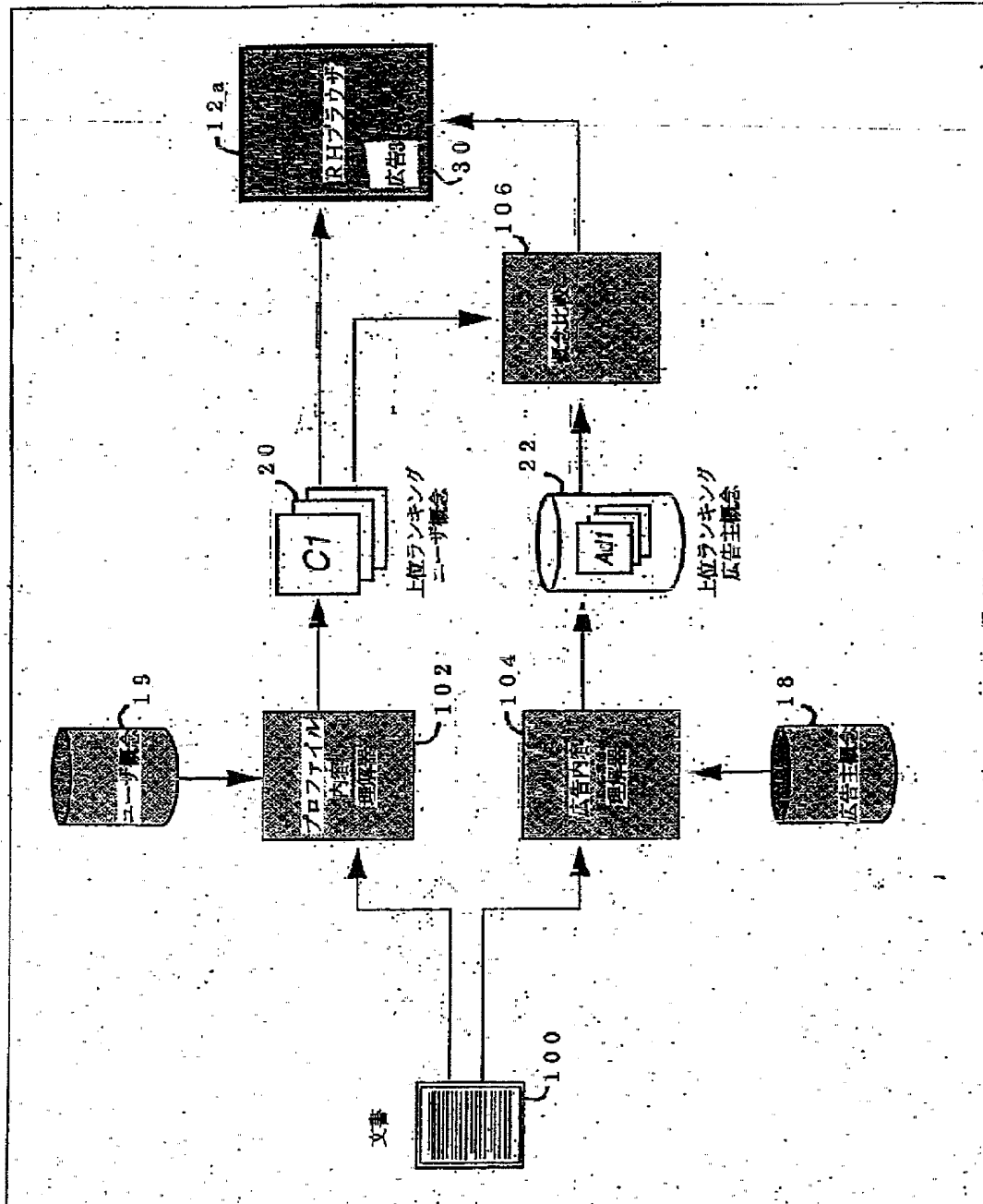
【図 1 A】

本発明の具体的な一実施例を実現する代表的なシステムの構成図



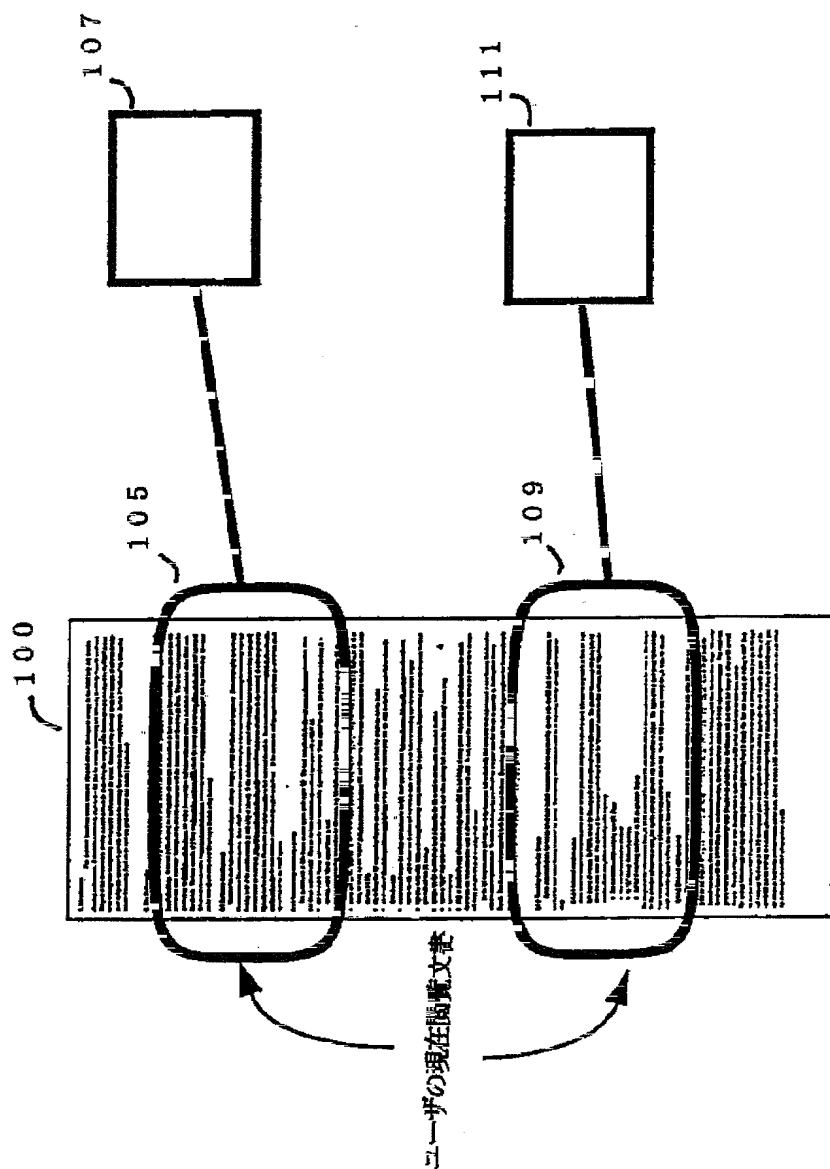
【図1B】

本発明の具体的な一実施例を実現する代表的なシステムの構成図



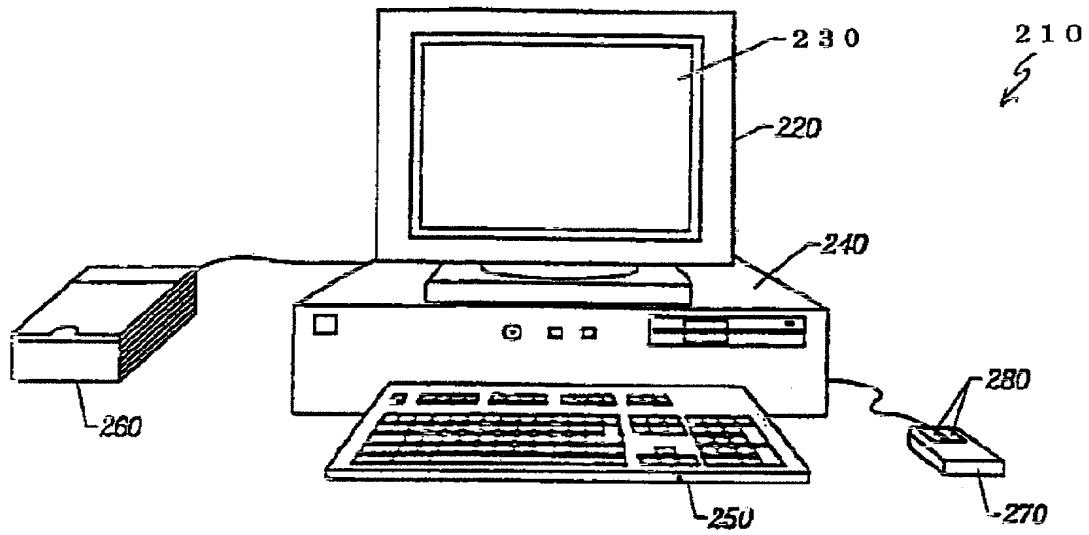
【図1C】

本発明の具体的な一実施例を実現する代表的なシステムの構成図



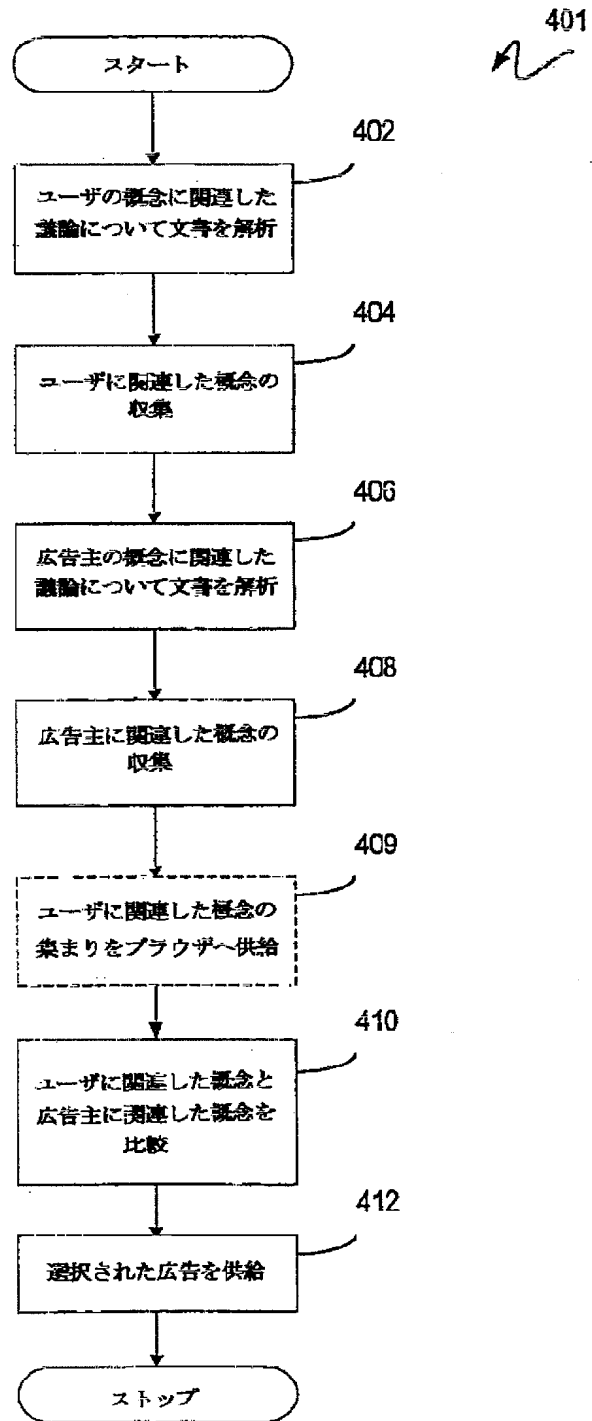
【図2】

本発明の具体的な一実施例におけるコンピュータシステムの概要図



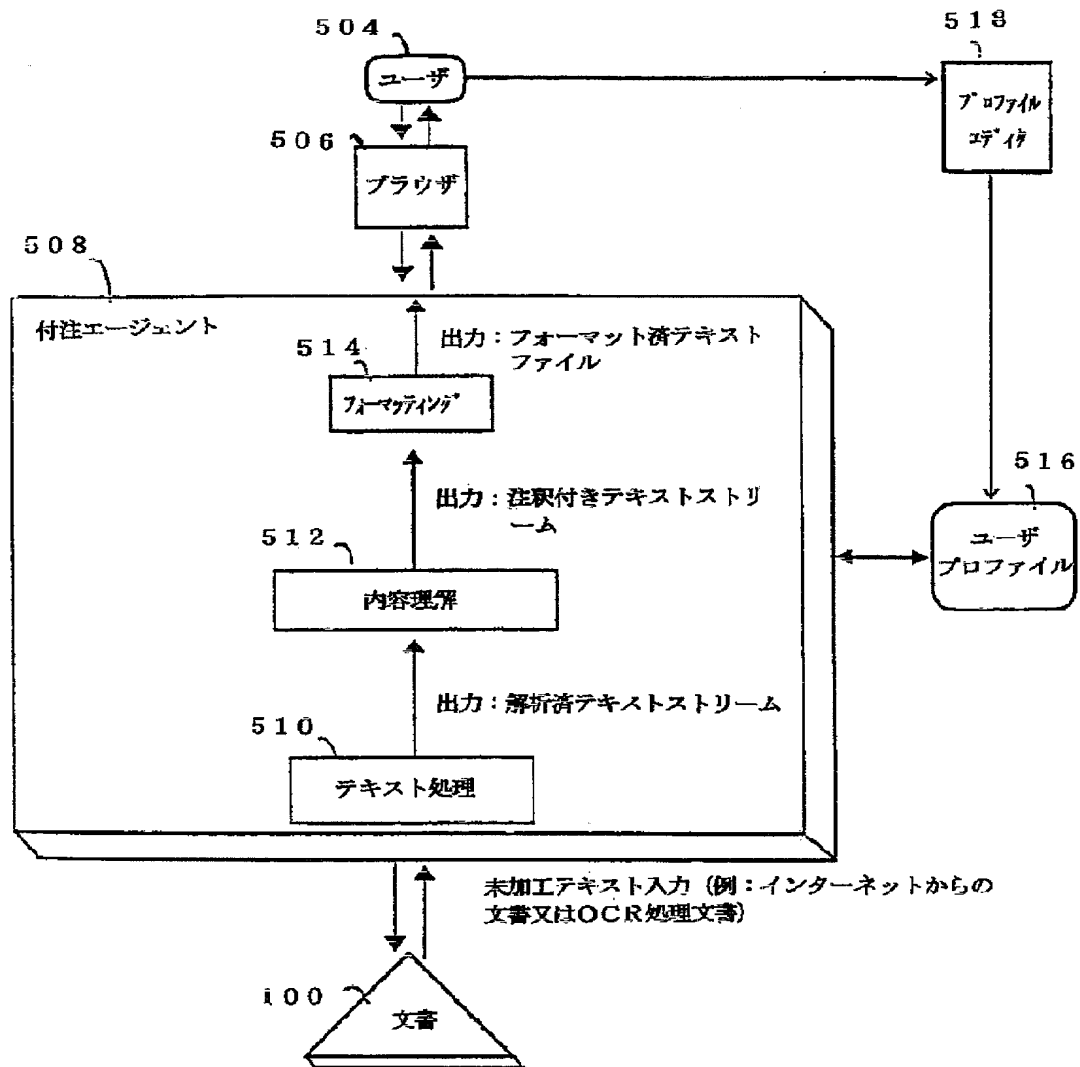
【図4A】

本発明の具体的な実施例における処理ステップの略フローチャート



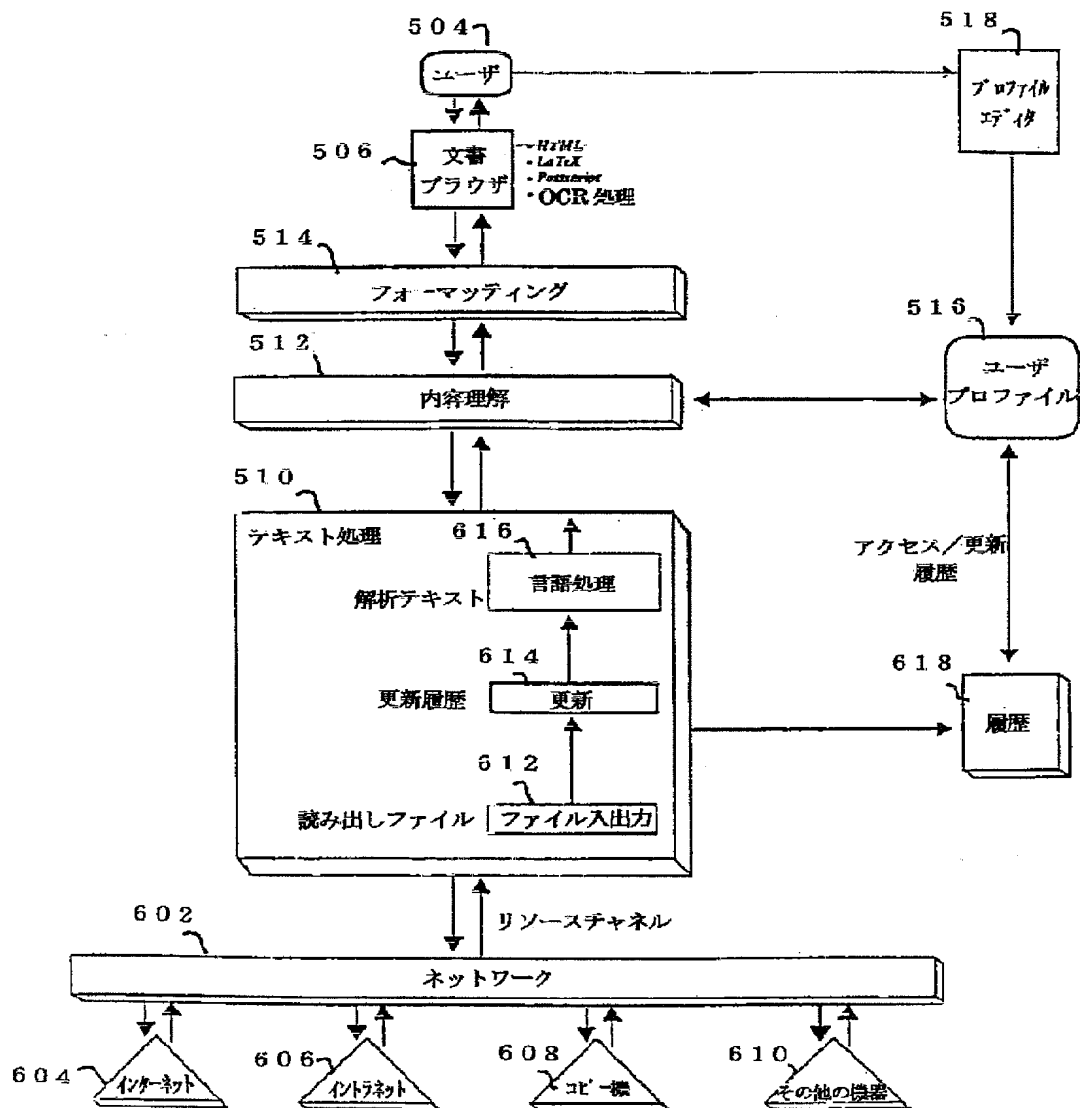
【図5】

本発明の具体的な一実施例における自動注釈付け処理の略ブロック図



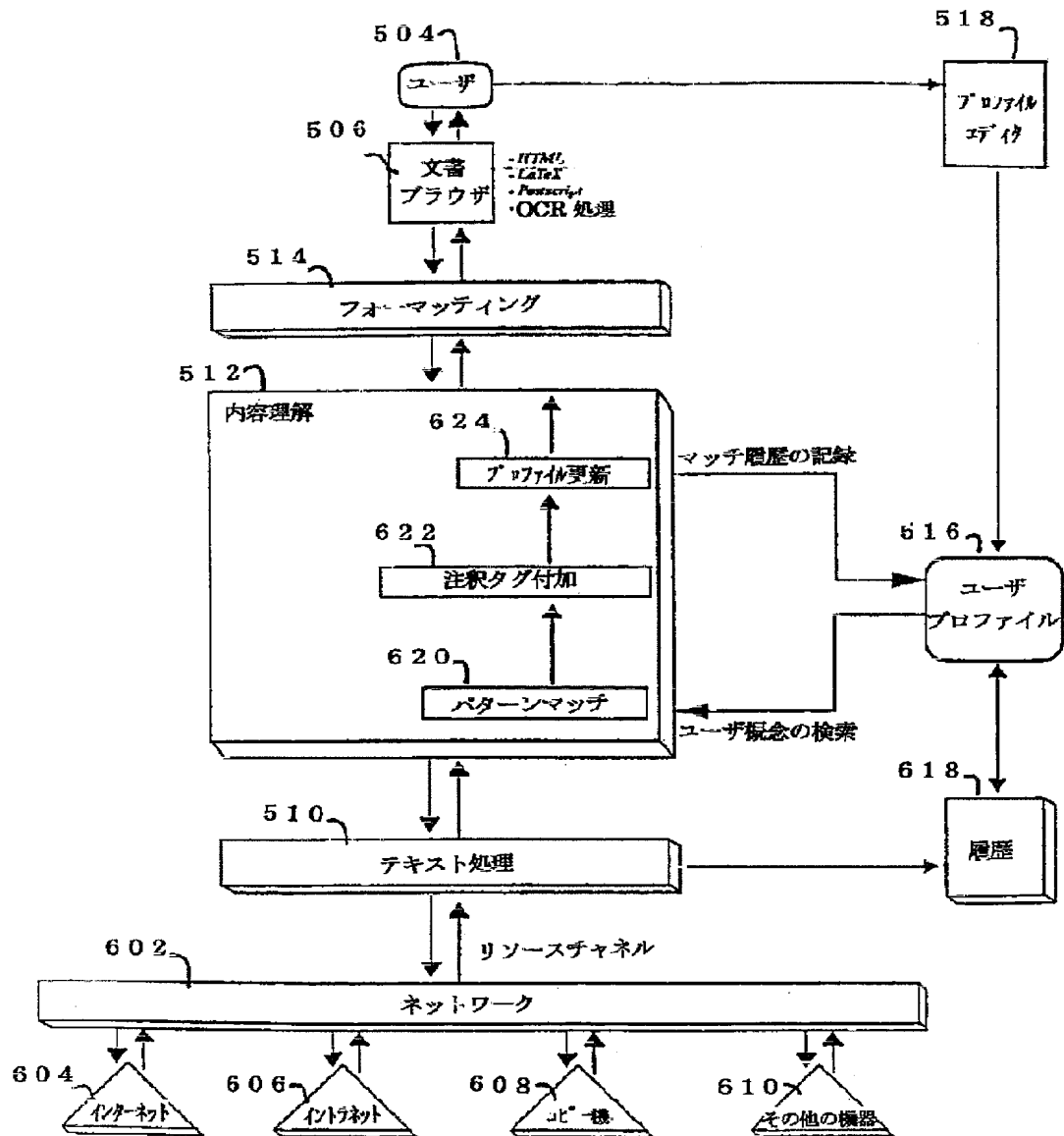
【図6A】

本発明の具体的な一実施例における自動注釈付け処理の略ブロック図



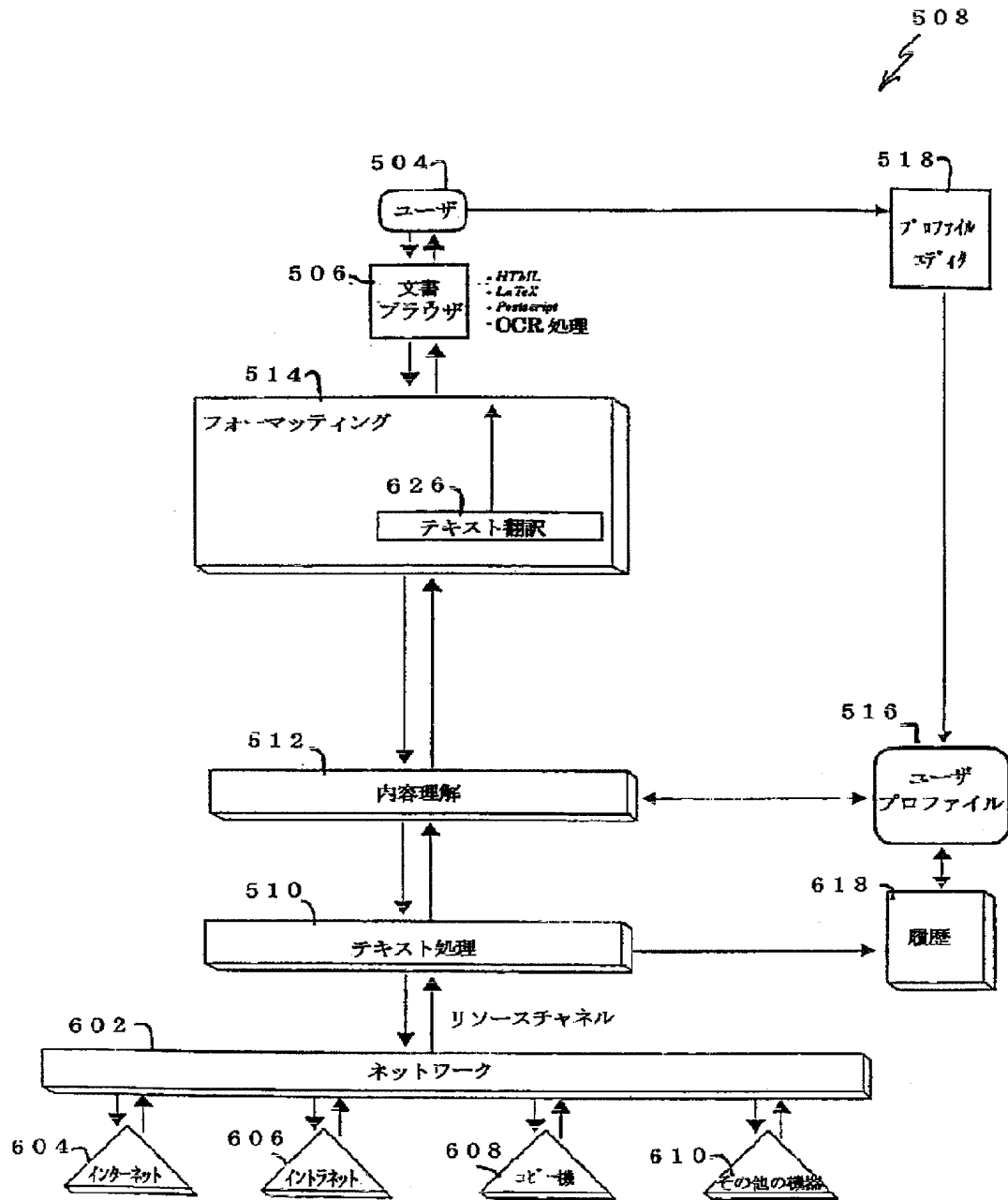
【図6B】

本発明の具体的な一実施例における自動注釈付け処理の略ブロック図



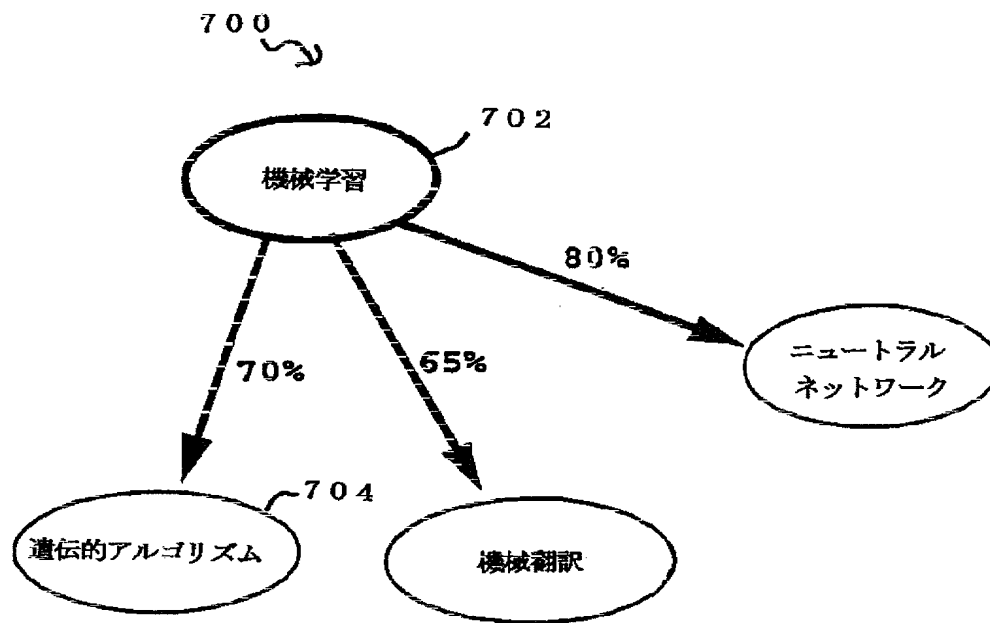
【図6C】

本発明の具体的な一実施例における自動注釈付け処理の略ブロック図



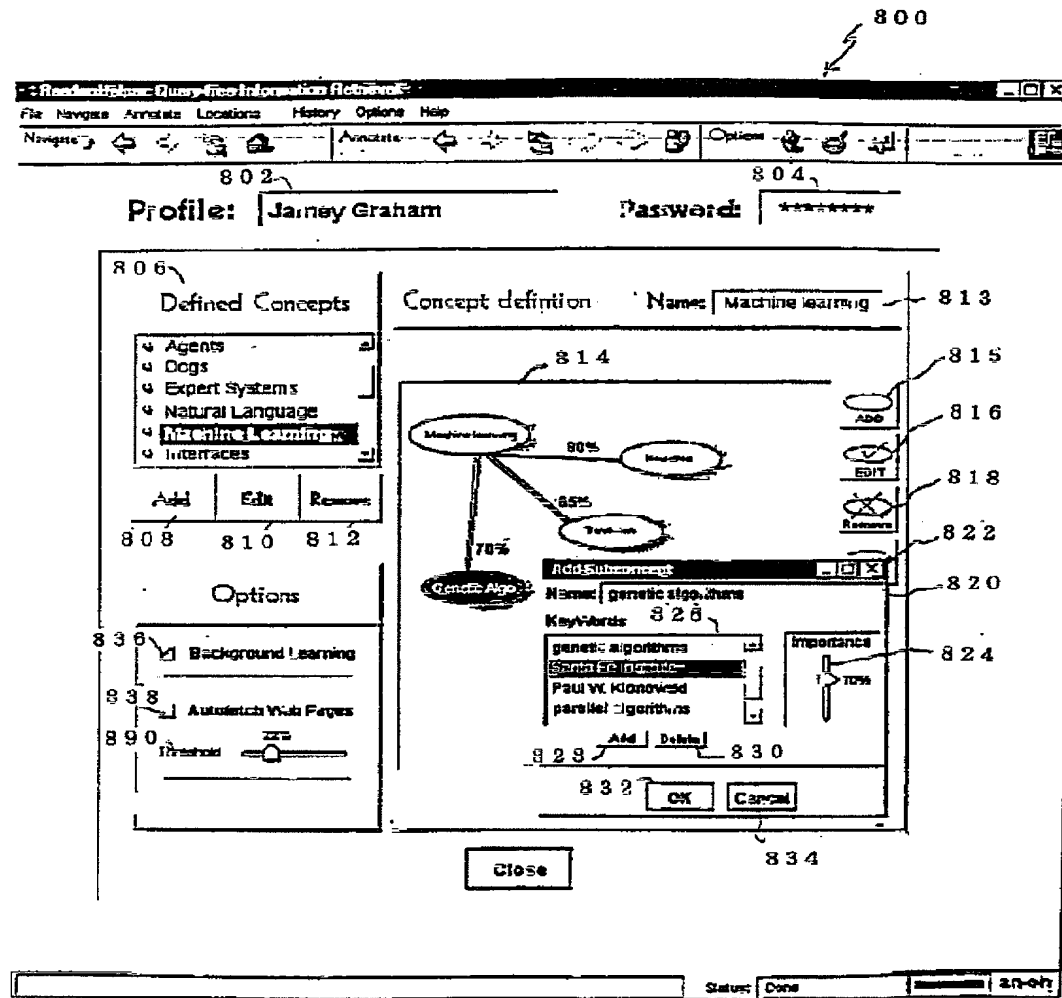
【図7】

本発明の具体的な一実施例における自動注釈付け処理で利用される代表的なベイズ信頼ネットワークの概略図



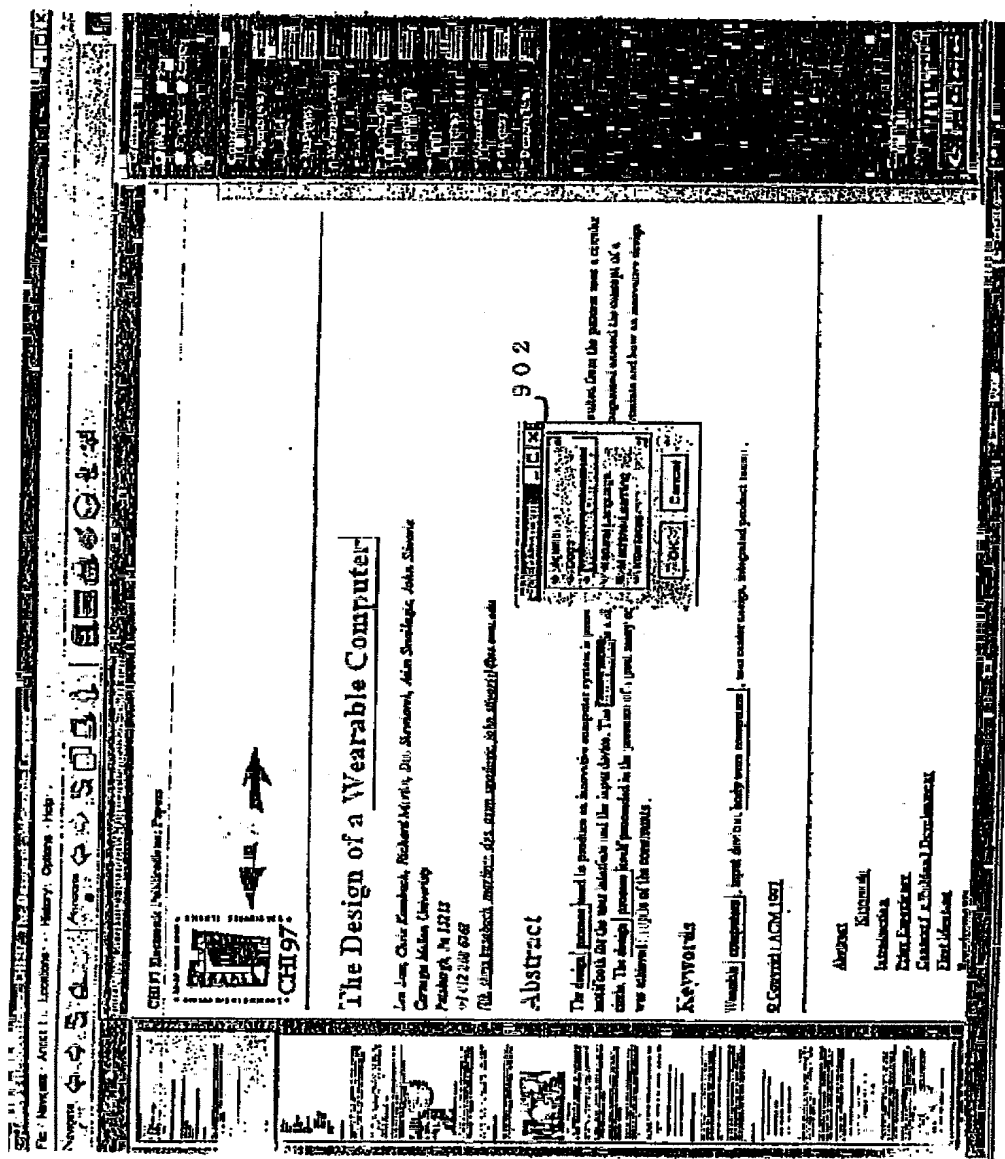
【図8】

本発明の具体的な実施例においてユーザプロフィールを定義するための代表的なユーザ・インタフェースの説明図



【図 9 A】

本発明の具体的な一実施例においてユーザ・フィードバックを行う代表的なインタフェースの説明図



【図10】

本発明の具体的な一実施例において処理される代表的なHTML文書の一部の説明図

```

1006 a      1006 b      1008 a      1008 b
We have approached this challenge by introducing an
<SPAN class=rhopic_Agent>Intelligent agent </SPAN> that analyzes interactions
between user and<SPAN class=rhopic_ExpertSystem>expert system </SPAN> and
automatically constructs database queries based on this analysis. The user is
unobtrusively notified when information relevant to the current diagnostic
context has been returned, and may immediately access it if desired. From the
user's perspective all database machinery is entirely transparent; indeed no
formal query language is even made available. Hence we term this approach
query-free information retrieval. <p>

1006 c      1008 c
As we hope will be apparent from what follows, the introduction of the
<SPAN class=rhopic_Agent>Intelligent agent </SPAN> additionally offers one
solution to a fundamental problem facing designers of cooperative information
systems: How can legacy systems of substantial complexity be integrated within
a larger system context? By requiring that all interactions with the legacy
database be mediated by the agent, we have been able to isolate the database
system cleanly while still supporting query-free information retrieval. <p>

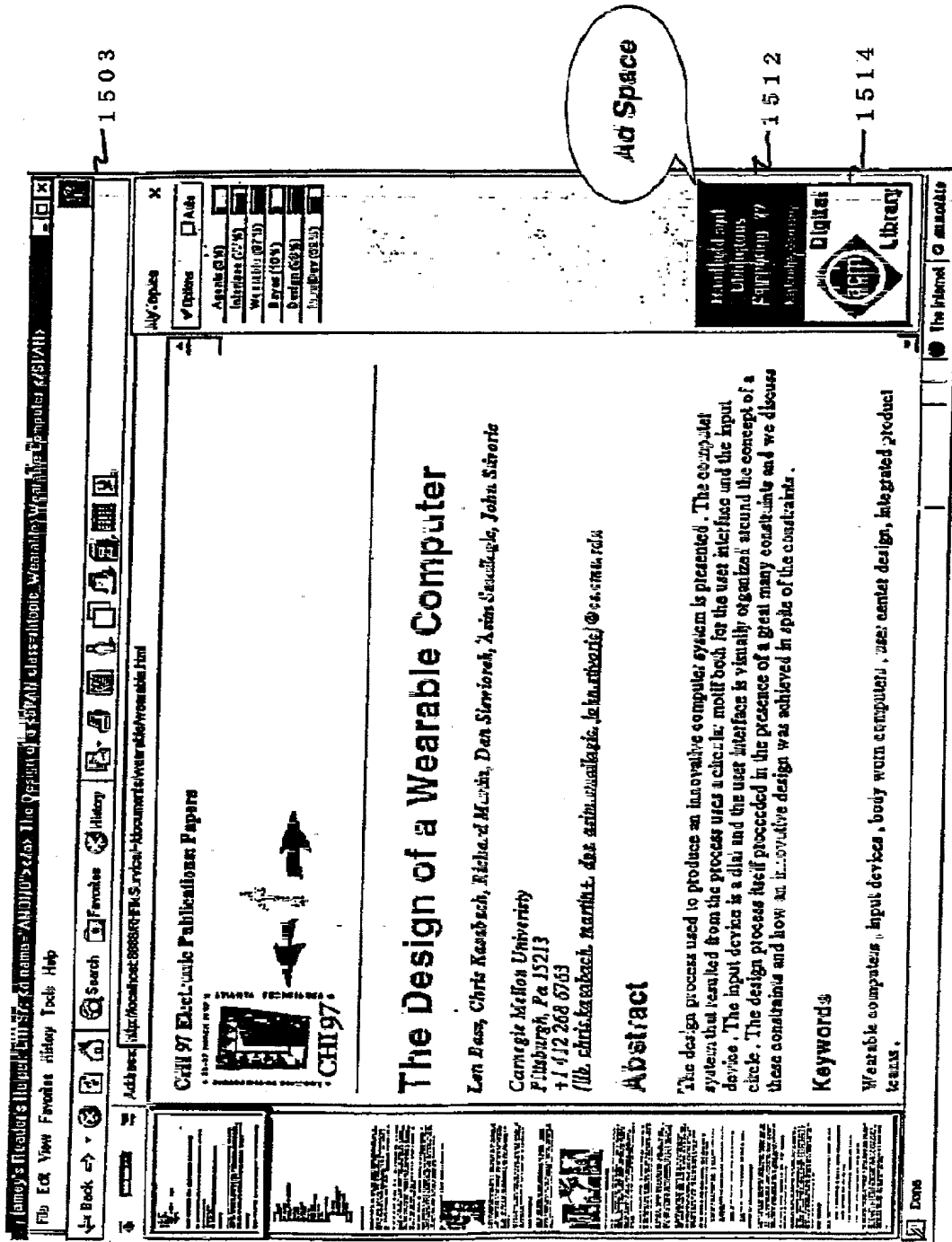
1006 d      1006 e      1008 d      1008 e
FIXIT is comprised of the three subsystems already mentioned: the probabilistic
<SPAN class=rhopic_ExpertSystem>expert system </SPAN>, the legacy full-text
database system (to which we added a new, semantically-based, indexing struc-
tures), and the <SPAN class=rhopic_Map>natural language </SPAN>
queries)., and the <SPAN class=rhopic_Agent>Intelligent agent </SPAN> that
effectively integrates them. The following sections describe these system com-
ponents, provide implementation details, illustrate the runtime behavior of
FIXIT, report on operational experience, and close with some observations about
query-free information retrieval and the potential for generalizing the under-
lying paradigm.<p>

1006 f      1008 f      1006 g      1008 g
<h2> FIXIT's System Components</h2>
We first describe the probabilistic <SPAN class=rhopic_ExpertSystem>expert
sub-system</SPAN> and the information retrieval sub-system. Before briefly
describing these, we stress that our purpose was not necessarily to advance the
capabilities of the individual components or indeed even to exploit fully the
best current technology; instead, we focus on their integration.<p>

```

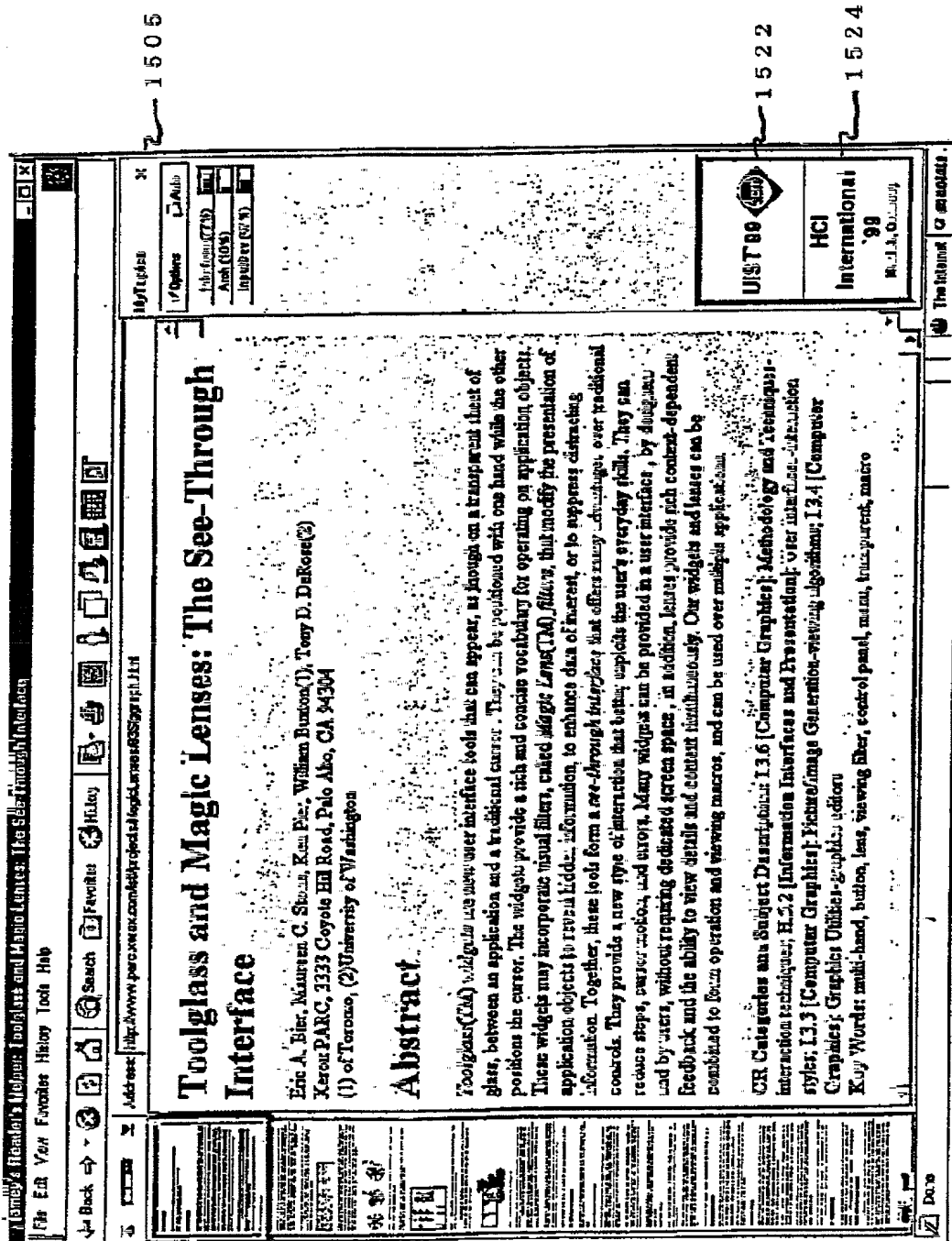
【図11A】

本発明の具体的な一実施例における代表的な画面表示の説明図



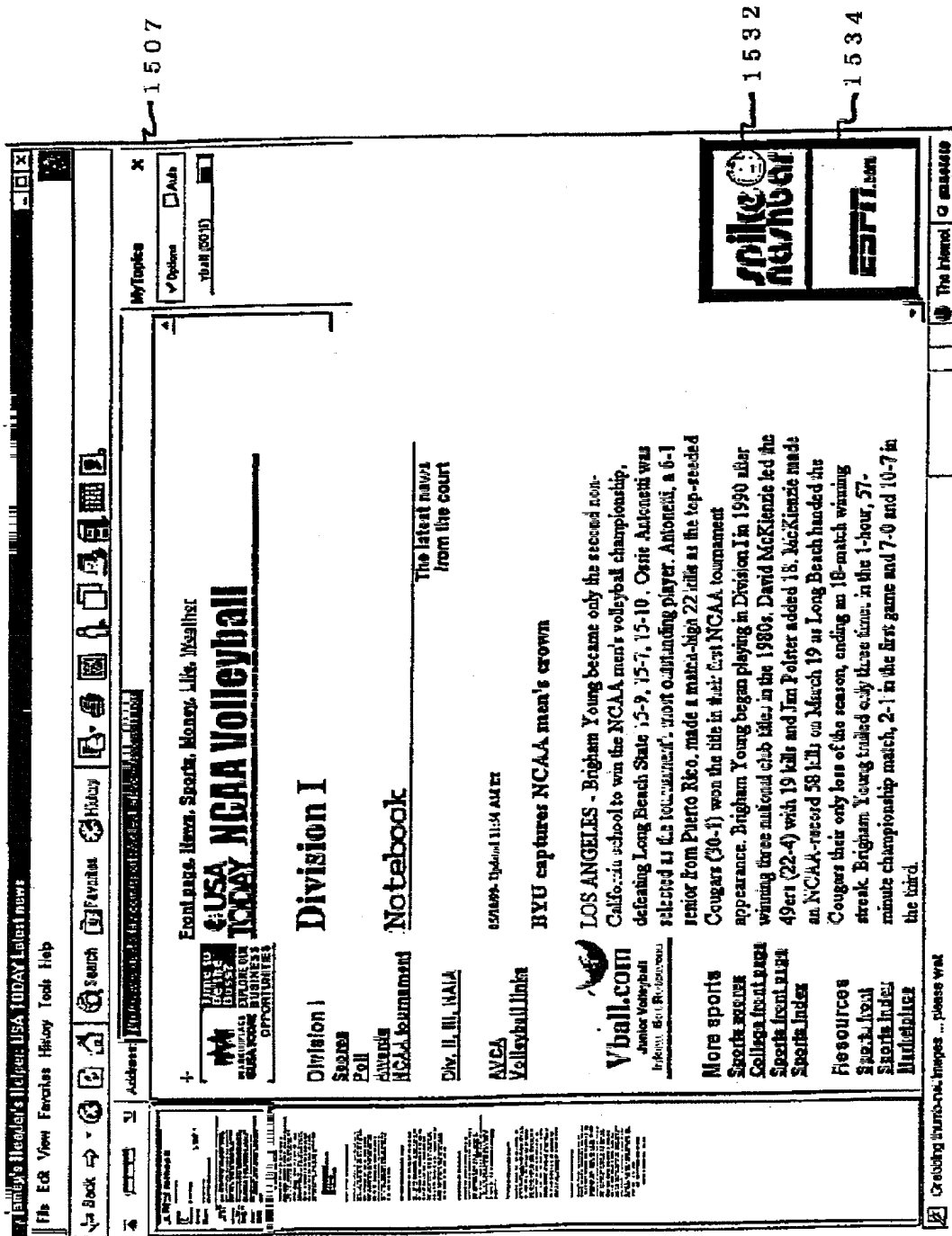
【図11B】

本発明の具体的な一実施例における代表的な画面表示の説明図

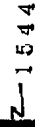


【図11C】

本発明の具体的な一実施例における代表的な画面表示の説明図

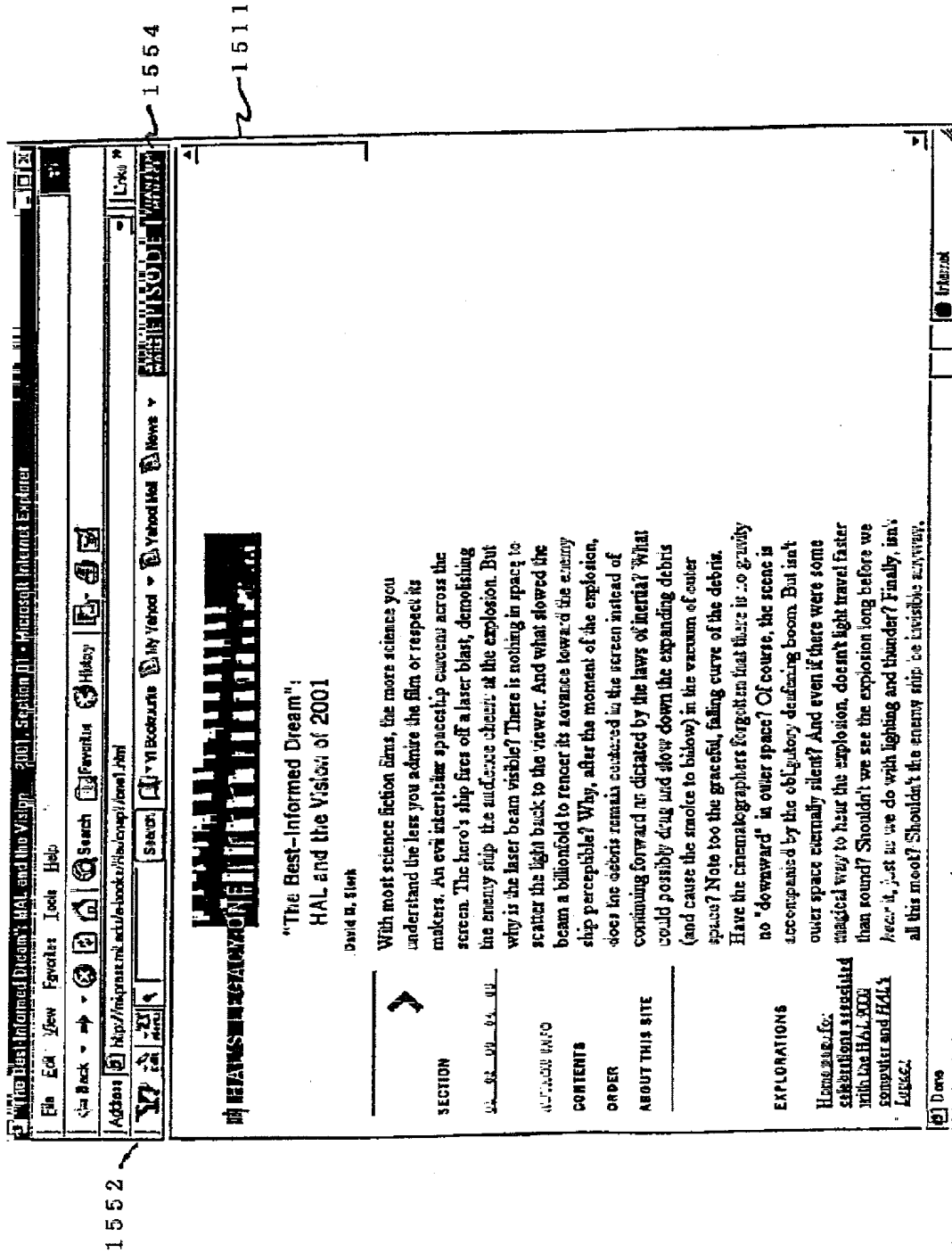


2-1509



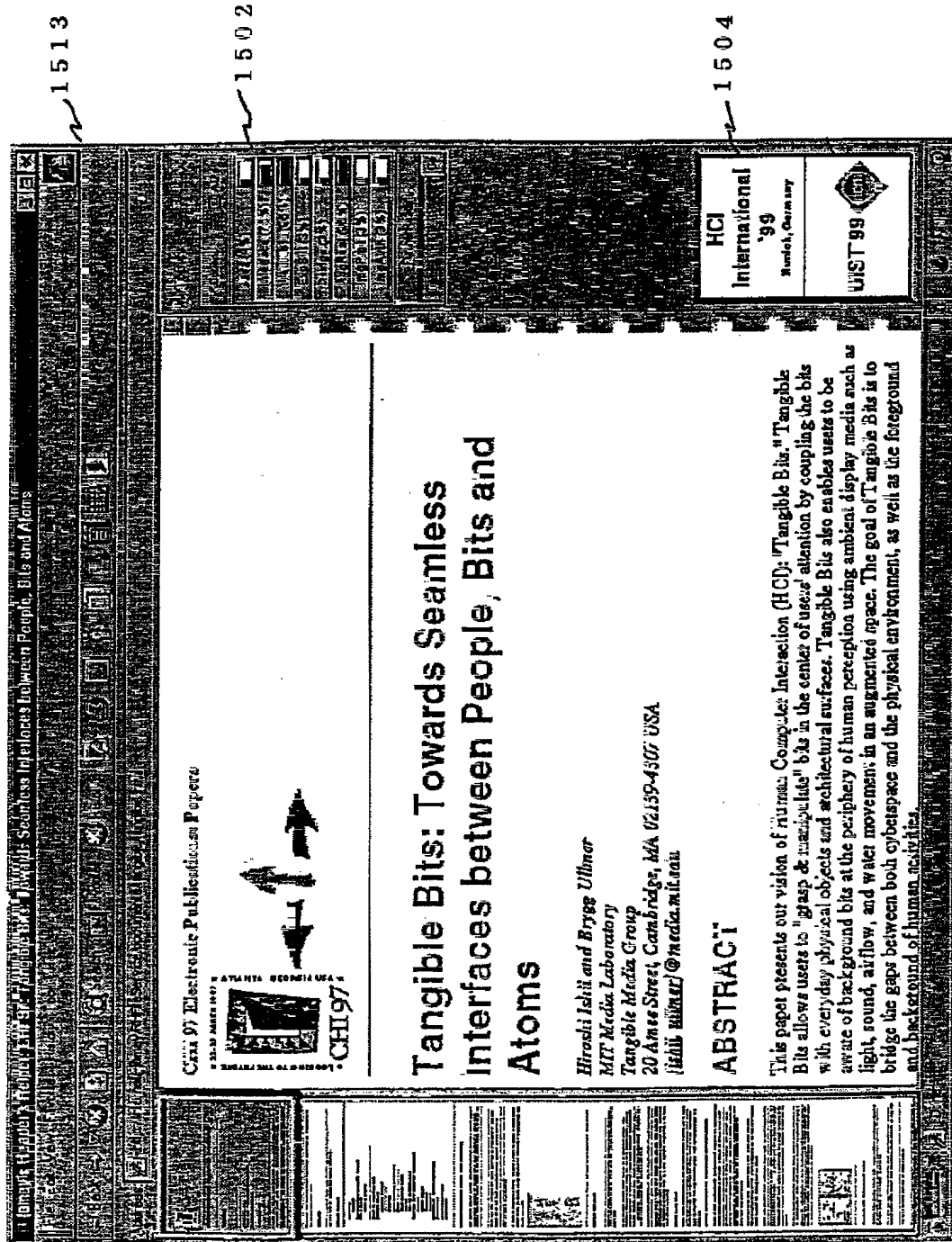
【図11E】

本発明の具体的な一実施例における代表的な画面表示の説明図



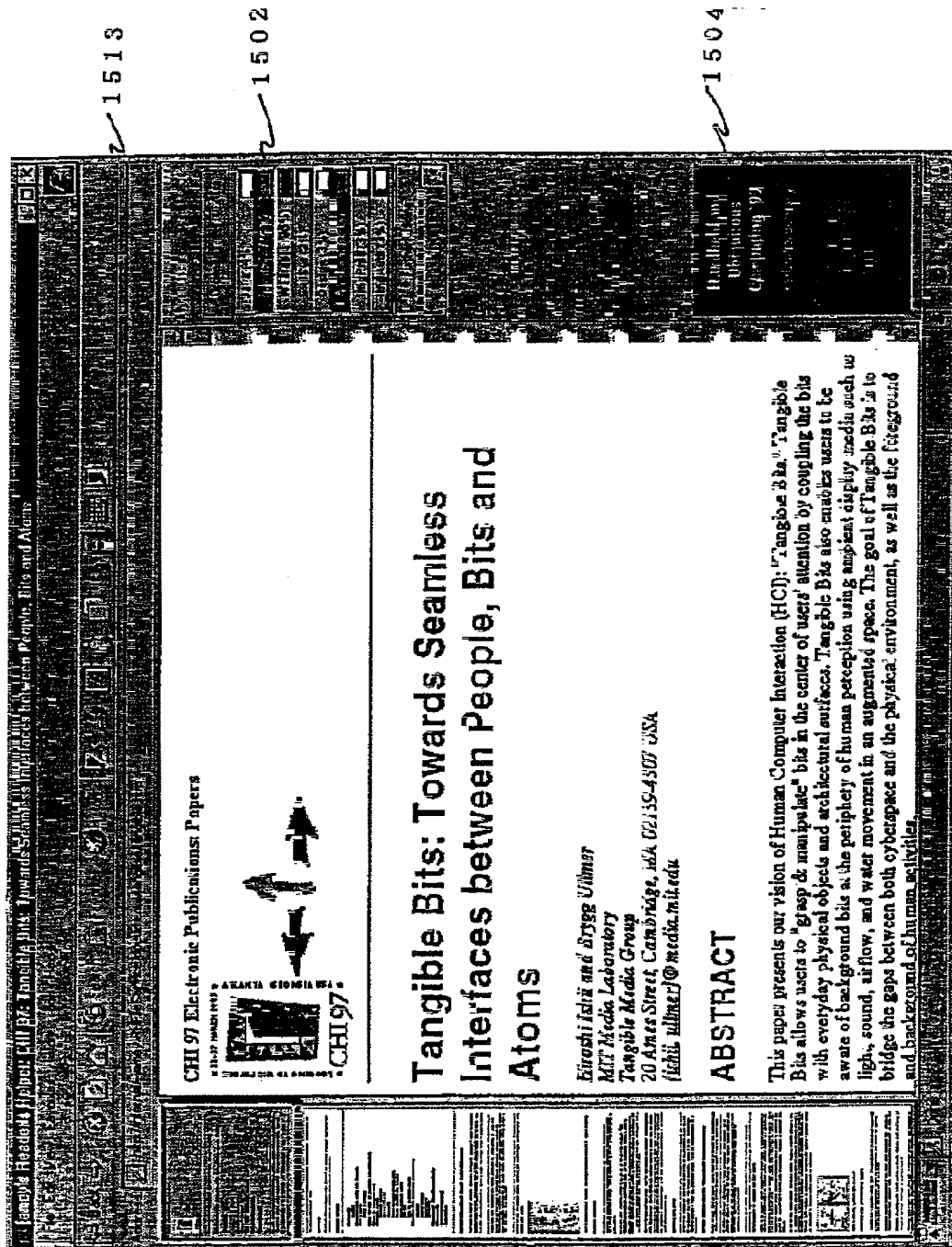
【図11F】

本発明の具体的な実施例による概念選択及び広告的強化処理の動作説明図



【図11G】

本発明の具体的な実施例による概念選択及び広告的確化処理の動作説明図



フロントページの続き

(72)発明者 デイビッド ジー ストーク
アメリカ合衆国, カリフォルニア 94025,
メンロ・パーク, サンド・ヒル・ロード
2882番, スイート115, リコーコーポレ
ーション内

Fターム(参考) 5B049 BB49 EE05 EE31
5B075 ND20 NK32 PR06